UNCOVERING HIDDDEN MICROBIAL METABOLISM USING 2D NMR- AND LC-MS-BASED COMPARATIVE METABOLOMICS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Joshua Andrew Baccile

January 2017

UNCOVERING HIDDDEN MICROBIAL METABOLISM USING 2D NMR- AND LC-MS-BASED COMPARATIVE METABOLOMICS

Joshua Andrew Baccile, Ph.D.

Cornell University 2017

Recent advances in genomic sequencing technology have facilitated relatively inexpensive, high-fidelity sequencing of microbial genomes. Increased genome sequencing capacity has been accompanied by developments in bioinformatics, which enable rapid homology searching to predict gene products, such as enzymes involved in the production of small-molecules. The recent surge in genomic information revealed that only a small percentage of genes involved in small-molecule biosynthesis had previously been characterized. In fact, it appears that we have only scratched the surface in understanding the full biosynthetic potential of microorganisms. The current void in small-molecule discovery is in part because of the particular challenges associated with identifying new compounds and in part because of the nature of microbial small-molecule production.

Perhaps the largest obstacle in the path to characterizing new small-molecules is the absence of a templated structure. For instance, in genomics, the number of building blocks is extremely limited, just four different bases, which fit together predictably through phosphodiester bonds. Moving up in complexity to proteomics, the number of building blocks increases to roughly twenty amino acids, however still assembled in a predictable manner, through peptide bonds. Additionally, protein primary sequence can be derived directly from the corresponding genetic sequence, which coupled with homology

searching enables quick judgments as to whether or not a detailed characterization is warranted. The situation in metabolomics is starkly more complicated as the number of building blocks increases to over one-hundred chemical entities that can be assembled in virtually any imaginable way that obeys the laws of chemistry. Moreover, small-molecule structural insights simply cannot be attained from the genetic sequence alone, meaning every compound must be individually characterized.

In addition to the challenges of small-molecule characterization, the very nature of microbial biosynthesis has significantly hindered more thorough annotation of the metabolome. Expression of genes encoding biosynthetic enzymes is tightly regulated and responsive to various natural environments. Simple laboratory culturing conditions usually fail to reproduce the necessary cues to promote small-molecule production. While media screening methods or so called fermentation optimization is an option to overcome this barrier, the power of genetic engineering offers a far more straightforward solution and has the added benefit of precision. The production of small-molecules can be turned on or eliminated in a controlled manner by creating over-expression or knock-out mutant microbial strains. Comparison of mutant and wild-type strains using differential analysis by 2D-NMR spectroscopy (DANS) and LC-MS-based comparative metabolomics enables a systems biology prospective of metabolic changes resulting from genetic engineering.

Described herein is the utilization of DANS and LC-MS-based comparative metabolomics to discover novel metabolites from orphan microbial biosynthetic gene clusters. Applied to the *has* gene cluster in *Aspergillus fumigatus* DANS revealed the biosynthesis of a novel iron(III)-complex, which was shown to increase *A. fumigatus'* virulence. Results from the *has* study led to an investigation of diketopiperazine formation

in the notorious gliotoxin biosynthetic pathway. Gliotoxin biosynthesis was shown to be dependent on a cluster-encoded cyclization domain, which ultimately functioned to cyclize a modified dipeptide into the diketopiperazine core of the gliotoxin structure. In another *A. fumigatus* gene cluster, called *fsq*, DANS and LC-MS-based comparative metabolomics revealed the elusive fungal isoquinoline formation pathway, which despite having been described in plants more than twenty years ago, remained unknown in fungi until now. The fungal isoquinoline pathway was shown to be conserved through the characterization of another isoquinoline producing gene cluster in the plant pathogenic fungus *A. flavus*. Lastly, the identification and characterization of an additional set of *fsq*-dependent metabolites, called the fumizinones, is described. The fumizinones derive from an alternative biosynthetic mechanism, resulting in pyrazinone formation, rather than isoquinoline formation, as observed for the main products of the *fsq* pathway.

BIOGRAPHICAL SKETCH

Joshua Andrew Baccile grew up in the small town of Big Flats, NY and graduated from Horseheads High School in June 2007. Immediately after high school Joshua enrolled at Corning Community College as a history major. Soon after beginning college Joshua became uninterested in history and decided that he would rather learn what had not yet happened, and thus, he embarked on the path of discovery by transferring to SUNY Cortland as a Chemistry major in the Fall of 2008. At SUNY Cortland Joshua conducted research on the synthesis of photocleavable peptide conjugates under the advisement of Professor Frank Rossi. In May 2011 Joshua completed his Bachelor of Science degree in Chemistry and matriculated at the Department of Chemistry and Chemical Biology, Cornell University to pursue his PhD. Joshua conducted his graduate research in the laboratory of Professor Frank C. Schroeder studying microbial secondary metabolism. During his time at Cornell Joshua was a Participant, Trainee, and Chair of the Chemical Biology Interface Training Program. In 2015 Joshua received the Simon Bauer Scholarship from the Department of Chemistry and Chemical Biology for his excellence in scientific research and communication.

Dedicated to my loving and supportive wife Laura.

ACKNOWLEDGMENTS

Special Committee Chair and Adviser: Professor Frank C. Schroeder.

Collaborating Principal Investigators: Professor Nancy P. Keller and Professor Dirk Hoffmeister.

Collaborating Researchers: Henry H. Le, Christian Gomez, Jason Hoki, Dr. Joseph E. Spraker, Brandon Phannonsteil, Dr. Jin Woo Bok, Dr. Philipp Weimen, and Dr. Wen-Bing Yin.

Special Committee: Professor Brian Crane and Professor Yimon Aye

NMR Facility Managers: Dr. Ivan Keresztes and David Kiemle

Organizations: Department of Chemistry and Chemical Biology and The Boyce Thompson Institute for Innovation and Discovery in the Life Sciences, Cornell University; Department of Medical Microbiology and Immunology, University of Wisconsin-Madison; Leibniz Institute for Natural Product Research and Infection Biology, Hans Knoll Institute; SUNY College of Environmental Science and Forestry, Analytical and Technical Services.

Funding: NIH Chemistry and Biology Interface Training Program Grant T32GM008500. TRIAD Foundation summer fellowship, and Frank C. Schroeder NIH R01GM112739-01.

TABLE OF CONTENTS

PREFACE

The purpose of this preface is to describe the overarching rationale for the research conducted towards this dissertation, to detail important background that pertains to all of the remaining chapters, and to provide clarity for field specific language used recurrently throughout each chapter. The general theme of this dissertation is using analytical chemistry tools to investigate and characterize metabolic pathways in filamentous fungi. While the work described here represents investigations specifically in fungal metabolism, the results are more broadly applicable to microorganisms, as fungi and bacteria (and other microbes) share related metabolic pathways[1].

The word metabolism refers to the broad set of chemical reactions that are necessary to sustain life[2]. Introduction of the qualifier “secondary” (metabolism), is to describe a subset of chemical reactions that are not essential for survival[3]. However, secondary metabolism is only unessential for survival in laboratory conditions, which represent unchallenging, nutrient rich conditions that would very rarely or altogether not occur in the wild[4]. Primary metabolism is then of course the set of chemical reactions that represent the bare-minimum necessary for survival in laboratory conditions[5]. The line between primary and secondary metabolism is blurred and as our understanding increases, the line becomes increasingly transparent. Eventually the community will likely not distinguish between primary and secondary metabolism at all, since, as has already been established to some extent, secondary metabolic pathways *are* in almost all circumstances necessary for survival in the wild[6]. The percentage of microbial genomes dedicated towards secondary metabolism is in itself a strong argument for the essential role of these complex chemical species[7]. For the sake of clarity, secondary metabolism

in reference to the research described in this dissertation pertains strictly to anabolic pathways that build larger molecules from the smallest possible building blocks, such as amino acids or acetyl-CoA's[8]. These anabolic pathways will henceforth be described as biosynthetic pathways or the biogenesis for a compound or a class of compounds.

Importance of studying secondary metabolism in filamentous fungi: Throughout our daily lives humans are continually in close-contact with fungi[9]. In particular, the fungus *Aspergillus fumigatus*, the focus of the work in this dissertation, is ubiquitously distributed in nature and its spores are regularly inhaled by humans (roughly 100-300 daily)[10]. The spores of *A. fumigatus* have evolved to be the perfect size such that they are inhaled and not exhaled out of the lungs[10,11]. Once inside the lungs, spores are normally engulfed by immune cells and destroyed, as is the case for a healthy individual, however, in an immunocompromised person, the immune system may fail to rid the lungs of all spores and the fungus will begin to grow[11]. The pulmonary fungal growth is then capable of spreading to other parts of the body, such as the brain, resulting in tissue necrosis and eventually death[11]. This infection, known as aspergillosis is extremely difficult to treat and accordingly is one of the leading causes of death for immunocompromised individuals in the United States[12]. In 2004 the Keller Lab at the University of Wisconsin-Madison showed that disruption of a global regulator of secondary metabolism, called LaeA, resulted in a loss of *A. fumigatus*' virulence in a mouse model, clearly demonstrating the importance of secondary metabolism for fungal survival during an infection[13]. In this dissertation unexplored LaeA regulated secondary metabolic pathways were targeted for characterization because of their potential importance for virulence. In addition to being

an important human pathogen, filamentous fungi also play a big role in the agricultural industry[14].

Parallel to the continuous direct interaction between *A. fumigatus* and humans is the interaction of plants and another related fungus, *A. flavus*, which thrives hand in hand with agricultural production[14]. *A. flavus* grows on a variety of different crops, such as corn, wheat, sorghum, and peanuts under high-humidity or stressful drought conditions. *A. flavus* became well known in the 1960's as the producer of Aflatoxin, which are potent carcinogens that poison feedstocks of nearly all types of livestock and some aquatic food sources, such as trout, in addition to food directly consumed by humans[15]. Across the globe losses due to Aflatoxin contamination are estimated to be upwards of $1-1.5 billion, which doesn't include the human health impact or crop yield losses[15]. *A. flavus* is just one of many plant pathogenic fungi that have the potential to cause high-value crop loss and health problems if not managed properly. Understanding secondary metabolism is essential to addressing both agricultural and health problems caused by fungi because of the direct implications of small-molecules (SMs) in their fitness and virulence. Additionally, annotation of the fungal metabolome will likely uncover new fundamental fungal biology.

While detrimental associations of filamentous fungi are notable, the contrary is equally as significant. Understanding SM production has the potential to yield new complex molecular architectures with unforeseen bioactivity. Historically speaking, a considerable percentage of therapeutics are derived from SMs of microbial origin[16]. A noteworthy class of drugs that are fungal natural products or derivatives thereof is the Statins, which have been used extensively to lower cholesterol[17]. The research described

in this dissertation is not specifically targeted at discovering or understanding the bioactivity of fungal metabolites. However, annotating a larger percentage of the currently scant metabolome inherently increases the chances of finding new molecules with medicinal importance. Furthermore, the central approach in this body of work (DANS-based comparative metabolomics) is directed toward the discovery of thermodynamically unstabile molecules, which arguably have the greatest potential for bioactivity, since potency and life time are inversely related in the context of SM interactions.

Natural history of microorganisms: A less widely discussed aspect of the rationale for studying microbial metabolism is its role in ecology. To draw an analogy to animals, metabolic products are microbial language. The production, transport, and break down of SMs represents the major form of communication between different species of microbes and between host-microbe interactions[18]. It would be difficult to understand the written English language without first having a grasp of the alphabet. This is exactly analogous to the situation with our understanding of microorganisms, in order to understand how they behave we must first understand the molecules they make, break, and trade. In the past it was often enough to “grind and find” or simply randomly extract a bacterial, fungal, or plant growth, or perhaps an animal. A host of important discoveries were made by this method, and this laid the foundation for the natural products chemists of today. Now, however, with the power of genetic engineering we have the ability to probe specific metabolic pathways. This serves not only to “trick” an organism into producing enough material for isolation and characterization, but also aids in the identification of regulatory elements and cross-pathway communication. Determining the underlying regulatory mechanisms for SM biosynthesis will be essential towards combating microbe-associated

diseases.

Microbial small-molecule chemistry: There is no strict definition for the term “small molecule”, however to the author of this dissertation it refers to any chemical species that has a molecular weight that is roughly 1500 amu or less. Of course, exceptions occur where a molecule of larger molecular weight still meets most criteria and would in fact be described as a SM. Microorganisms are the most prominent producers of SMs on the earth and are capable of assembling extremely complex molecular architectures, which has driven the field of synthetic organic chemistry and provided us with the vast majority of antibiotics[16].

Genes encoding enzymes responsible for the assembly of microbial secondary metabolites usually exist in proximity to each other in the genome, comprising a biosynthetic gene cluster (BGC)[19]. Gene clustering has been exploited as a mechanism to enable facile upregulation or deletion of entire biosynthetic pathways by a single promoter switch[8]. In bacteria gene clustering seems to be the absolute rule for biosynthesis (there may be exceptions), in fungi and other eukaryotes, however, there are occasionally genes in other locations of the genome that are still genuine components of the BGC-encoded biosynthetic pathway[20]. Gene clustering has also enabled more rapid compound discovery and helped eliminate rediscovery via bioinformatics analyses[20]. Algorithms have been designed to use homology searching to predict the locations of BGCs in the genome and to characterize the corresponding gene products. Overall, homology-based methods have been incredibly useful but occasionally they are misleading because despite sharing high homology, a pair of BGCs may produce completely different SM products. The simple definition of what constitutes a BGCs is the

existence of a gene encoding a core enzyme surrounded by one or many genes encoding tailoring enzymes (see following section)[21].

Microbial biosynthetic gene cluster composition: Most characterized microbial BGCs encode for core enzymes that fall into one of three major families, the terpene cyclases (TCs), the polyketide synthases (PKSs), and the non-ribosomal peptide synthetases (NRPSs)[19,22]. TCs function through the catalysis of cationic cyclization reactions of linear isoprene oligomers[22]. PKSs closely resemble standard fatty acid biosynthesis enzymes and function by promoting Claisen condensations between successive malonyl-CoA's[19]. NRPSs function through the assembly-line style of loading amino acids, catalyzing peptide bond formation, and then releasing peptide products[19,23]. Beyond the core enzyme, BGCs across the three major biosynthetic families include similar tailoring enzymes, which serve to modify the scaffold produced by the core enzyme. Some of the common tailoring enzymes include cytochrome p450 oxidases, *N*- and *O*-methyltransferases, and prenyltransferases[24]. BGCs also often include transporter proteins, such as the hallmark multi-drug resistant ABC-transporters[25]. Importantly, many BGCs have one or more transcription factors that regulate the expression of all cluster genes (the usual target for genetic engineering to facilitate whole-cluster upregulation). Nearly all of the research described in this dissertation is about NRPS-based BGCs.

While it is outside the scope of this preface to exhaustively cover NRPS background it is useful to describe some of the fundamentals aspects of NRPS chemistry. The first step in NRPS biosynthesis is the activation of the free carboxy group of an amino acid by the adenylation (A) domain. Following A domain-dependent activation the amino-adenylate ester is captured by a free thiol of the phosphopantetheinyl arm on the thiolation

(T) domain. Then, a condensation (C) domain catalyzes the peptide bond formation between two T domain bound thioesters, resulting in a nonribosomal peptide (NRP). The NRP can be modified further by additional domains present in the NRPS, or by tailoring enzymes, or can be simply released from the NRPS as an unmodified peptide. The mechanism and timing of release dictates the biosynthetic scheme for a particular BGC. For example, if an NRP is released largely unmodified, then ensuing modifications by tailoring enzymes will occur by shuttling around the NRP from one enzyme to the next until it has been completely elaborated. If, on the other hand, the modifications occur while the peptide is still bound to the NRPS, then the released NRPS product is the final pathway product and the substrate of each tailoring enzyme is an NRPS-tethered intermediate. Distinguishing between these two modes of NRPS biosynthesis is important for understanding the true biosynthetic pathway of many microbial metabolites.

2D-NMR and LC-MS-based comparative metabolomics powered by genetic engineering: Previous methods for the detailed analysis of BGCs involved *in vitro* reconstitution of individual cluster enzymes to determine their function[26]. This was however, hopelessly laborious, as many enzymes are extremely difficult to express and purify in functional form *in vitro*. A recent and major advance, and the foundation of the research described in this dissertation was the development of differential analysis by 2D-NMR spectroscopy (DANS)-based comparative metabolomics[27]. The principle of DANS-based comparative metabolomics is to overlay 2D-NMR spectra of whole-metabolome extracts from two different mutant microbial strains[27]. The two mutants differ in that one is an overexpression (OE) strain and the other is a knock-out (KO) strain for the same BGC[28]. Thus, features present in the 2D-NMR spectrum of the metabolome from the OE

strain and absent in the corresponding KO spectrum are BGC-associated metabolites. Furthermore, structural features can be determined directly from the OE spectrum, which means in one step the identification of cluster associated metabolites as well as partial (occasionally full) structures can be attained[29,30]. Coupling the DANS-based approach to liquid chromatography-mass spectrometry (LC-MS) based comparison of mutant strains provides another dimension to determine differential features corresponding to cluster associated metabolites and is necessary to determine molecular formulae. Utilizing this approach for the analysis of mutant strains for individual cluster genes allows for the determination of a "snap-shot" of the biosynthetic pathway. The work described in this dissertation, utilizing this comparative metabolomics approach provides detailed information about the biosynthetic function of entire BGCs and individual enzymes with previously unachievable expediency and thoroughness.

Preview of chapters: Each chapter in this dissertation represents a separate but related exploration of secondary metabolic gene clusters in the pathogenic fungi, *A. fumigatus* or *A. flavus*. Chapters 1 and 3 are composed of reformatted and updated co-first author manuscripts written by the author of this dissertation. Additional text and figures have been included where appropriate to provide more details pertaining to data acquisition and analysis, hypothesis formation, and open questions. Chapters 2 and 4-5 consist of separate, unpublished research projects that have been summarized specifically for this dissertation. Supporting information for Chapters 1-5 can be found in Appendices A-E. Importantly, all of the projects represented in this dissertation were performed in collaboration with several other labs. Author contributions for each chapter will be described here and not within each chapter.

Chapter 1, A Nonribosomal Peptide Synthetase-Derived Iron(III) Complex from the Pathogenic Fungus *Aspergillus fumigatus*[28]: Chapter 1 describes the characterization of a previously orphan gene cluster, which we called the *has* cluster. We choose to explore the *has* cluster because it was both positively regulated by host-exposure and the global regulator essential for virulence, LaeA. Comparison of overexpression and knock-out strains of *hasA,* the cluster transcription factor, using DANS and LC-MS-based comparative metabolomics revealed the final product of the pathway to be an iron(III)-complex, hexadehydroastechrome (HAS, **Figure P1**). Additional DANS-based analyses of double mutant strains where tailoring genes were deleted in cluster overexpression background revealed several novel intermediates in the biosynthetic pathway, allowing us to propose a model for HAS biosynthesis. Deletion of the *has* cluster NRPS, HasD, led to a significant upregulation in the fumitremorgin gene cluster, revealing cluster cross-talk within *A. fumigatus*.

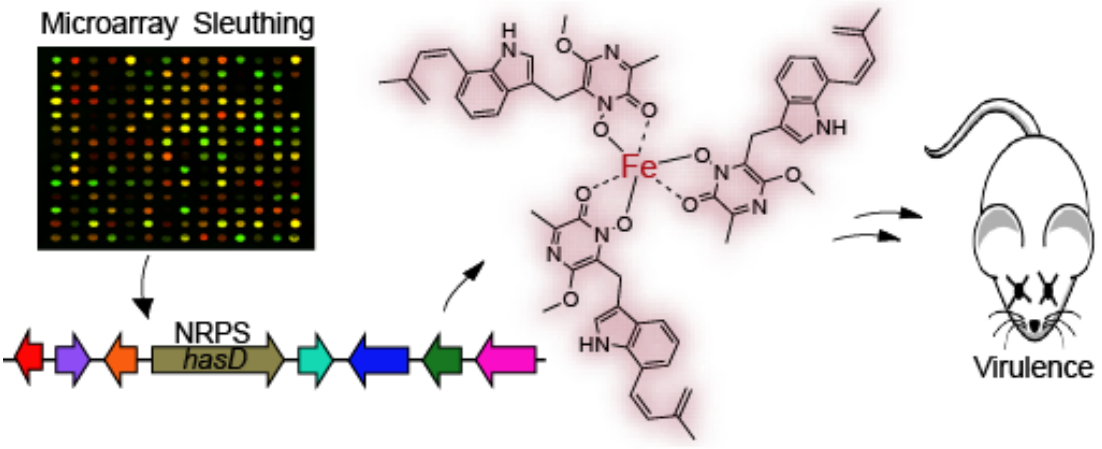


Figure P1. Overview of Chapter 1. Microarray sleuthing revealed *has* cluster expression increased under virulence conditions, which upon targeted overexpression and deletion enabled DANS-based discovery and characterization of the final product as the novel iron(III)-complex. Overexpression of the *has* cluster created a super pathogenic strain of *A fumigatus* that increased mortality in mice with pulmonary infections.

Authors' contributions: Joshua A. Baccile isolated and characterized all compounds, constructed the figures, and wrote the manuscript. Wen-Bing Yin created and cultured all *A. fumigatus* strains and wrote the manuscript. Jin Woo Bok and Yiming Chen performed microarray sleuthing experiments. Nancy P. Keller and Frank C. Schroeder designed the research and wrote the manuscript.

Chapter 2, Second condensation-like domain in GliP is essential for diketopiperazine cyclization in gliotoxin biosynthesis: Chapter 2 revisits the well-studied and notorious gliotoxin biosynthetic pathway in *A. fumigatus*. The core enzyme in the *gli* cluster, GliP, consist of seven domains, A_1-T_1-C_1-A_2-T_2-C_2-T_3. The terminal C_2-T_3 domains have long been thought not to participate in the biosynthesis of gliotoxin. However, LC-MS-based comparative metabolomics of GliP truncation mutant *A. fumigatus* strains revealed the C_2 domain is necessary for gliotoxin production *in vivo* and functions as a cyclization domain (C_T), rather than a condensation domain (**Figure P2**).

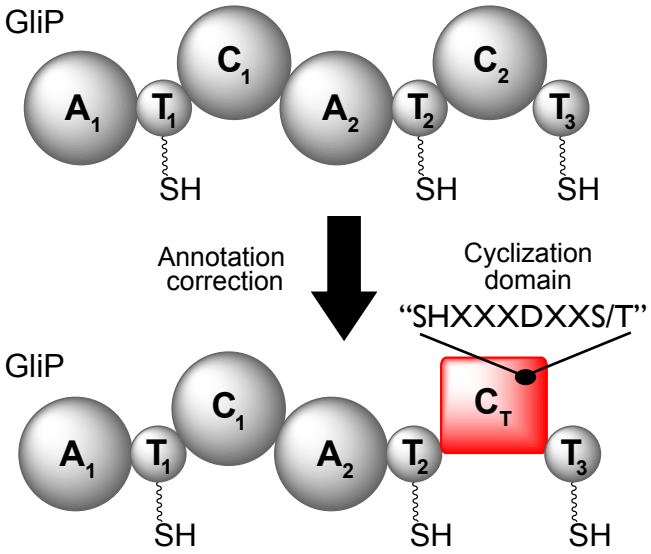


Figure P2. Overview of Chapter 2. The gli cluster NRPS, GliP, was previously annotated as having two condensation domains. We show the second condensation-like domain is actually a cyclization domain (C_T).

In vitro reconstitution of the GliP truncation mutant proteins and the terminal C_2-T_3 standalone domains confirmed the cyclization function of the C_2 domain. The results presented in Chapter 2 support a *gli* pathway biosynthetic model where tailoring modifications occur on a T_3-tethered linear dipeptide, which is then cyclized off by the C_2 (now C_T) domain.

Authors' contributions: Joshua A. Baccile designed the research, performed LC-MS data collection and analysis, and wrote the manuscript. Henry H. Le performed *in vitro* GliP assays. Brandon Phannenstiel and Jin Woo Bok created mutant *A. fumigatus* strains. Christian Gomez assisted with LC-MS data collection. Eileen Brandenburger and Dirk Hoffmeister conducted A domain activation assays. Nancy P. Keller and Frank C. Schroeder helped design the research.

Chapter 3, Plant-like biosynthesis of isoquinoline alkaloids in *Aspergillus fumigatus*[31]: Chapter 3 focuses on one member of a large but neglected family of orphan BGCs featuring truncated non-ribosomal peptide synthetases, that, because of their truncated nature, cannot produce canonical peptides. Such non-canonical NRPSs are widespread among fungi, however, their biosynthetic roles have remained obscure. We show that the *fsq* biosynthetic pathway produces a novel family of isoquinoline alkaloids, the fumisoquins, which are produced via a multi-step enzymatic cascade that is entirely analogous to the biosynthesis of an important class of plant alkaloids, e.g. in poppy (**Figure P3**) Additionally, the fumisoquin pathway provided evidence for a new carbon-carbon bond formation mechanism via cooperative action of the truncated NRPS and an ATP-grasp enzyme.

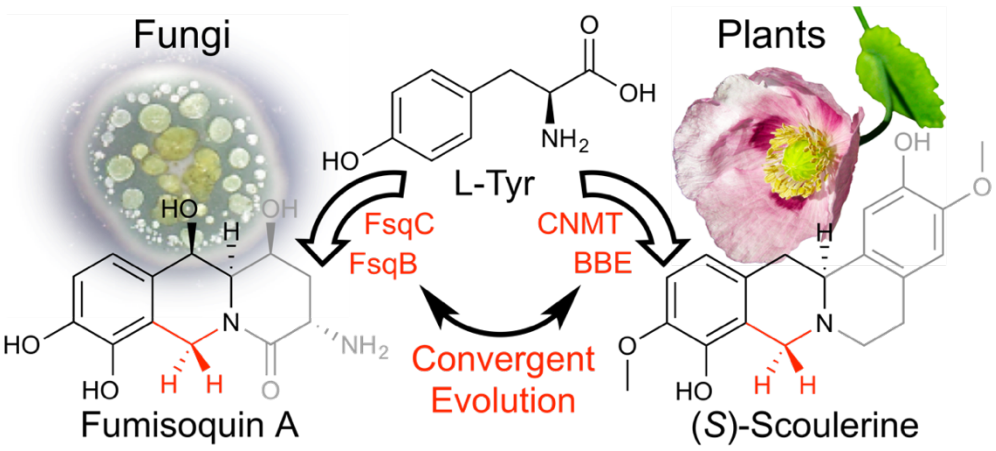


Figure P3. Overview of Chapter 3. Fungi and plants separately evolved analogous sets of enzyme to incorporate methylene group (red) that forms the isoquinoline present in Fumisoquins and Scoulerine.

Authors' contributions: Joshua A. Baccile and Frank C. Schroeder identified and characterized metabolites and biosynthetic pathways. Joseph E. Spraker and Jin Woo Bok created A. fumigatus mutants. Joshua A. Baccile and Christian Gomez performed stable-isotope labeling experiments. Henry H. Le, Eileen Brandenburger, and Dirk Hoffmeister contributed biochemical assays. Juliane Macheleidt and Axel A. Brakhage contributed to the identification of new metabolites. Joshua A. Baccile, Joseph E. Spraker, Dirk Hoffmeister, Nancy P. Keller and Frank C. Schroeder wrote the manuscript.

Chapter 4, Biosynthesis of the structurally unprecedented imizoquins in the plant pathogenic fungus *Aspergillus flavus*: In connection with Chapter 3, Chapter 4 is the exploration of another isoquinoline producing BGC, called the *imq* cluster, however, in *A. flavus* instead of *A. fumigatus*. LC-MS comparative metabolomics followed by detailed 2D NMR spectroscopic analysis of overexpression strains of the transcription factor ImqK revealed *imq*-dependent production of a class of tripeptide-derived alkaloids, the imizoquins (**Figure P4**). These compound include the expected isoquinoline moiety as

part of a tricyclic imidazo[2,1-*a*]isoquinoline ring system, for which there are no other examples among known natural products. As in the case of the fumisoquins, the isoquinoline portion of the imizoquins appears to be derived from hydroxylation, *N*-methylation and subsequent oxidative cyclization of tyrosine.

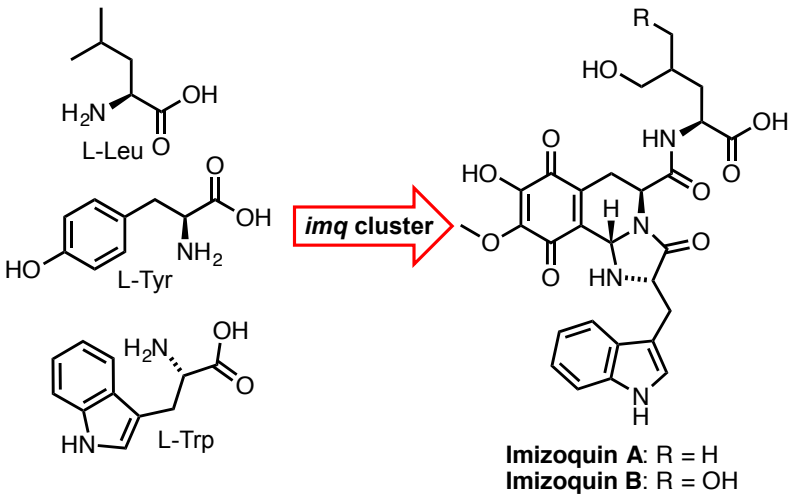


Figure P4. Overview of Chapter 4. The *imq* cluster incorporates L-Leu, L-Tyr, and L-Trp to produce the Imizoquins, novel tricyclic imidazo[2,1-a]isoquinolin alkaloids from the plant pathogenic fungus *A. flavus*.

Authors' contributions: Joshua A. Baccile identified and characterized metabolites. Joseph E. Spraker created *A. flavus* mutant strains. Saima Khalid performed total antioxidant capacity assays. Christian Gomez cultures *A. flavus* and helped in preparative compound isolation. Nancy P. Keller and Frank C. Schroeder helped design the research.

Chapter 5, Biosynthesis of the fumizinones: secondary products of the *fsq* cluster in *Aspergillus fumigatus*: During the characterization of the *fsq* cluster described in detail in Chapter 3, we discovered that an additional set of metabolites was *fsq*-dependent, which were not isoquinoline alkaloids like the fumisoquins. These additional

fsq-dependent metabolites were instead found to be of the pyrazinone class of natural products, so we called them the fumizinones (**Figure P5**). Stable-isotope amino acid feeding experiments allowed us to determine that the fumizinones were derived from L-leucine and L-serine, and their formation was likely dependent on a unique carbon-carbon bond forming reaction catalyzed by the NRPS-like enzyme FsqF.

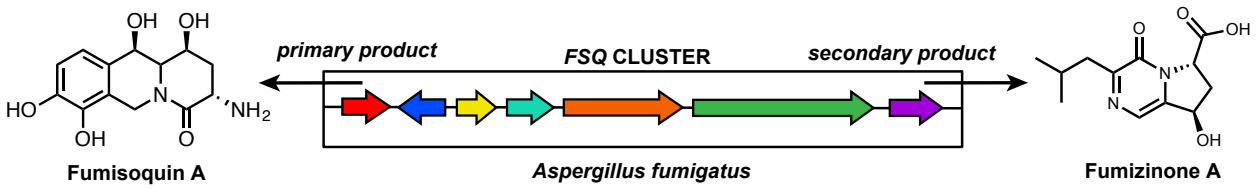


Figure P5. Overview of Chapter 5. The *fsq* cluster in *A. fumigatus* produces two distinct classes of molecules, the fumisoquins (left) and the fumizinones (right).

Authors' contributions: Joshua A. Baccile and Frank C. Schroeder identified and characterized metabolites and biosynthetic pathways. Joseph E. Spraker and Jin Woo Bok created *A. fumigatus* mutants. Joshua A. Baccile and Christian Gomez performed stable-isotope labeling experiments. Nancy P. Keller and Frank C. Schroeder helped design the research.

REFERENCES

1. Luckner, M. *Secondary Metabolism in Microorganisms, Plants and Animals*. (2013).

2. White, A., Handler, P. & Smith, E. *Principles of Biochemistry*. (1959).

3. Keller, N. P., Turner, G. & Bennett, J. W. Fungal secondary metabolism—from biochemistry to genomics. *Nat Rev Microbiol* **3**, 937-947, (2005).

4. Yu, J. H. & Keller, N. Regulation of secondary metabolism in filamentous fungi. *Annu Rev Phytopathol* **43**, 437-458, (2005).

5. Drew, S. W. & Demain, A. L. Effect of primary metabolites on secondary metabolism. *Annu Rev Microbiol* **31**, 343-356, (1977).

6. Stadler, M. & Keller, N. P. Paradigm shifts in fungal secondary metabolite research. *Mycol Res* **112**, 127-130, (2008).

7. Challis, G. L. Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiol* **154**, 1555-1569, (2008).

8. Yin, W. & Keller, N. P. Transcriptional regulatory elements in fungal secondary metabolism. *J Microbiol* **49,** 329–339, (2011).

9. Taylor, L. H., Latham, S. M. & Mark, E. J. Risk factors for human disease emergence. *Philos Trans R Soc Lond B Biol Sci* **356**, 983-989, (2001).

10. Latgé, J. P. Aspergillus fumigatus and aspergillosis. *Clin Microbiol Rev* **12**, 310-350, (1999).

11. Denning, D. W. Invasive aspergillosis. *Clin Infect Dis* **26**, 781-803, (1998).

12. Stevens, D. A., Graybill, J. R. & Group, P. A. S. Invasive Aspergillosis Disease Spectrum, Treatment Practices, and Outcomes. *Medicine* **79**, 250-260, (2000).

13. Bok, J. W. & Keller, N. P. LaeA, a regulator of secondary metabolism in Aspergillus spp. *Eukaryotic Cell* **3,** 527–535, (2004).

14. Amaike, S. & Keller, N. P. Aspergillus flavus. *Annu Rev Phytopathol* **49,** 107–133, 2011).

15. KLICH, M. A. Aspergillus flavus: the major producer of aflatoxin. *Mol Plant Pathol*

8, 713–722, (2007).

16. Newman, D. J. & Cragg, G. M. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J Nat Prod* **75,** 311–335, (2012).

17. Unit, E. S. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *Lancet* **366**, 1267-1278, (2005).

18. Nutzmann, H.-W. *et al.* Bacteria-induced natural product formation in the fungus Aspergillus nidulans requires Saga/Ada-mediated histone acetylation. *Proc Natl Acad Sci USA* **108,** 14282–14287, (2011).

19. Fischbach, M. A. & Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev* **106,** 3468–3496, (2006).

20. Yadav, G., Gokhale, R. S. & Mohanty, D. Towards Prediction of Metabolic Products of Polyketide Synthases: An In Silico Analysis. *PLoS Comput Biol* **5,** e1000351, (2009).

21. Walsh, C. T. The chemical versatility of natural-product assembly lines. *Acc Chem Res* **41,** 4–10, (2008).

22. Greenhagen, B. & Chappell, J. Molecular scaffolds for chemical wizardry: learning nature's rules for terpene cyclases. *Proc Natl Acad Sci USA* **98**, 13479-13481, (2001).

23. Du, L. & Lou, L. PKS and NRPS release mechanisms. *Nat Prod Rep* **27,** 255-278, (2010).

24. Walsh, C. T., Chen, H., Keating, T. A. & Hubbard, B. K. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr Opin Chem Biol* **5**, 525-534, (2001).

25. Higgins, C. F. ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **8**, 67-113, (1992).

26. Sattely, E. S., Fischbach, M. A. & Walsh, C. T. Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. *Nat Prod Rep* **25**, 757-793, (2008).

27. Pungaliya, C. *et al.* A shortcut to identifying small molecule signals that regulate behavior and development in Caenorhabditis elegans. *Proc Natl Acad Sci USA* **106,** 7708–7713, (2009).

28. Yin, W.-B. *et al.* A Nonribosomal Peptide Synthetase-Derived Iron(III) Complex from the Pathogenic Fungus Aspergillus fumigatus. *J Am Chem Soc* **135,** 2064–2067, (2013).

29. Butcher, R. A. *et al.* The identification of bacillaene, the product of the PksX megacomplex in Bacillus subtilis. *Proc Natl Acad Sci USA* **104,** 1506–1509, (2007).

30. Forseth, R. R. & Schroeder, F. C. NMR-spectroscopic analysis of mixtures: from structure to function. Curr Opin Chem Biol **15**, 38-47, (2010).

31. Baccile, J. A. *et al.* Plant-like biosynthesis of isoquinoline alkaloids in Aspergillus fumigatus. *Nat Chem Biol* **12**, 419-424, (2016).

CHAPTER 1

A NONRIBOSOMAL PEPTIDE SYNTHETASE-DERIVED IRON(III) COMPLEX FROM THE PATHOGENIC FUNGUS *ASPERGILLUS FUMIGATUS*

Abstract: Small molecules (SMs) play central roles as virulence factors of pathogenic fungi and bacteria; however, genomic analyses suggest that the majority of microbial SMs have remained uncharacterized. Based on microarray analysis followed by comparative metabolomics of overexpression/knockout mutants we identified a tryptophan-derived iron(III)-complex, hexadehydro-astechrome (HAS), as the major product of the cryptic *has* non-ribosomal peptide synthetase (NRPS) gene cluster in the human pathogen *Aspergillus fumigatus*. Activation of the *has* cluster created a highly virulent *A. fumigatus* strain that increased mortality of infected mice. Comparative metabolomics of different mutant strains allowed to propose a pathway for HAS biosynthesis and further revealed cross-talk with another NRPS pathway producing the anti-cancer fumitremorgins.

Introduction: *Aspergillus fumigatus* is a ubiquitous filamentous fungus and causative agent of invasive aspergillosis (IA), a life-threatening disease. IA associated mortality ranges from 50% to 90%.[1] One of the few proteins that have been shown to be required for *A. fumigatus'* pathogenicity is LaeA, a conserved virulence factor in filamentous fungi.[2-5] LaeA is a global regulator of secondary metabolism in fungi and effects virulence, in part, by positive regulation of multiple small-molecule gene clusters.[3-6] For example, LaeA controls the biosynthesis of the epidithiodiketopiperazine gliotoxin,[7-10] which contributes to *A. fumigatus'* success as a pathogen.[11] We reasoned that SMs most likely to have a role in virulence would be positively regulated by both *laeA* and exposure to host environments (host exposure itself or hypoxic environments).

Identification of the *has* biosynthetic gene cluster and predicted gene products: By comparing microarrays of genes down-regulated in *laeA* deletants but up-regulated by host exposure/hypoxia,[6,12-14] we identified an eight-gene SM cluster named *has* meeting the aforementioned criteria. The eight cluster genes encode putative C6 transcription factors (HasA and HasF), one transporter (HasB), one *O*-methyltransferase (HasC), one non-ribosomal peptide synthetase (NRPS, HasD, previously denoted as Nrps 5[15] and PesF[16]), one 7-dimethylallyltryptophan synthase (DMATS)[17] (HasE), one putative FAD binding protein (HasG) and one cytochrome P450 (HasH) (**Figure 1.1a** and **Tables A1-3**). Often SM clusters are regulated by C6 transcription factors. Because *hasA* was the 3rd most down-regulated gene in the Δ*laeA* array and strongly induced in conidia during neutrophil exposure,[6,12-14] we chose to overexpress (OE) this gene to facilitate chemical characterization of *has*-associated metabolites.[18,19] Replacement of the *hasA* promoter

with the constitutive *gpdA* promoter (OE::*hasA*) resulted in induction of the other seven genes (**Figure 1.1b**) and yielded a pink pigmented strain (**Figure A1a**). Three cluster flanking genes, (AFUB_036320, AFUB_036310, and AFUB_036220, **Figure 1.1a**, **b**) were not regulated by *hasA* and thus likely defined the boundaries of the cluster.

Genetic engineering and comparative metabolomics reveal *has* products: To probe the functions of the individual *has* genes, we additionally constructed knock-out mutants for *hasB-hasH* in the OE::*hasA* background (**Figure A2**). Previous biochemical characterization of the DMATS HasE had proven its capacity to prenylate a variety of indole derivatives in position 7, with L-tryptophan as a preferred substrate;[17] however, no cluster-associated metabolites were identified. We employed HPLC-MS (**Figure 1.1c**) and NMR-based comparative metabolomics[8,19,20] to probe for *has*-dependent metabolites in the OE::*hasA* strain. Analysis of the ^{1}H NMR spectrum from the OE::*hasA* extract revealed signals indicative of a 3-substituted indole moiety that were absent in the WT extract spectrum (**Figure A3**). All signals of this compound suffered from intense line broadening, suggestive of a paramagnetic species such as an iron complex. Comparison

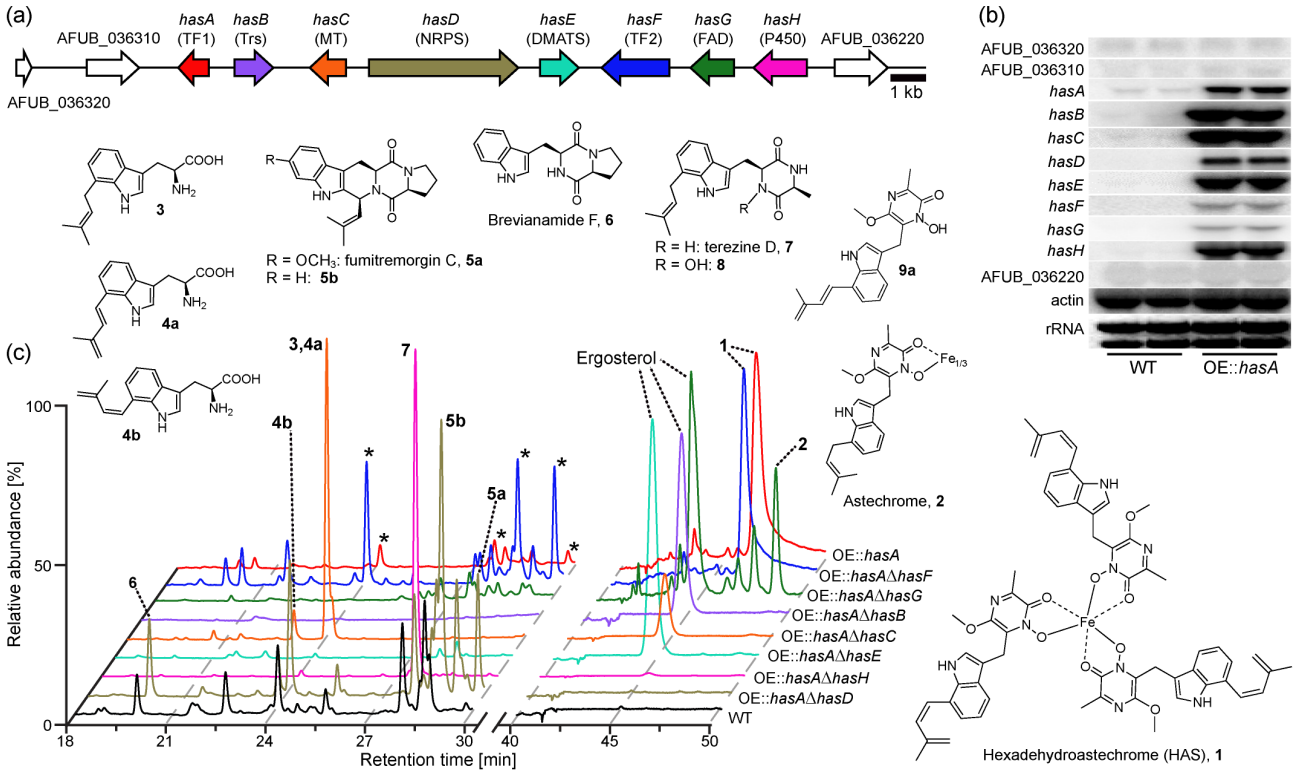


Figure 1.1: *hasA* controls *has* cluster gene expression and metabolite production. (a) Schematic of *has* gene cluster and putative assignments. *hasA* and *hasF*: C6 transcription factors; Trs: transporter; MT: methyltransferase; NRPS: non-ribosomal peptide synthetase; DMATS: dimethylallyltryptophan synthase; FAD: FAD binding domain protein; P450: cytochrome P450. Flanking genes are shown in white. (b) Northern analysis of *has* cluster gene expression in the WT and OE::*hasA* strains. (c) HPLC UV chromatograms for *A. fumigatus* mutant strains and structures of *has*-dependent metabolites; (*) denotes HAS decomposition products. Ergosterol amounts were highly variable between replicates and are not *has*-dependent.

of high-resolution (HR)-HPLC-MS data from the OE::*hasA* and WT extracts showed a OE::*hasA*-specific ion cluster eluting at 46.8 min whose *m/z* values and isotopic distribution suggested a molecular formula of $C_{60}H_{60}FeN_9NaO_9$ (*m/z* calcd. 1129.3761; obsvd. 1129.3751), suggesting an Fe(III) complex. Supporting the requirement of iron for production of this metabolite, OE::*hasA* cultures in iron-deficient medium did not produce the peak at 46.8 min or the pink color (**Figure. A1b**). Because extreme line-broadening prevented full spectroscopic characterization of the *has*-dependent iron complex, we

treated the isolated starkly pink compound with strong base to precipitate iron as the hydroxide. Following iron removal we obtained the *O*-methylated diketopiperazine derivative **9a**, whose structure was determined using a routine set of 2D NMR spectra and HR-MS (**Figure 1.1c**). Based on the combined spectroscopic data, the iron-containing metabolite was identified as a trimeric complex of **9** with Fe(III), similar to the previously described *A. terreus* metabolite, astechrome (1113.1, $[M+H]^+$).[21] Accordingly, we named the *has*-dependent metabolite we identified from *A. fumigatus* hexadehydroastechrome "HAS" (**1**; **Figure 1.1c**).

To gain insight into HAS biosynthesis, we examined the metabolomes of the seven double mutants carrying deletions of the other *has* genes in OE::*hasA* background (**Figure A2**) by 2D NMR and HPLC-UV-MS for the presence of pathway intermediates or shunt metabolites, focusing on compounds whose UV and NMR spectra suggested incorporation of tryptophan. Putative pathway-related metabolites were characterized by HR-MS and purified as needed to complete NMR-spectroscopic assignments (**Tables A4-9**). Loss of the second transcription factor (OE::*hasAΔhasF*) did not result in any detectable metabolomic change compared to the OE::*hasA* strain, indicating that under the growth conditions used in this study HasF does not affect the *has* pathway (**Figure 1.1c** and **Figure A2**). In contrast, deletion of the transporter (HasB) led to complete abolishment of HAS production and no intermediates or shunt metabolites were observed (**Figure 1.1c**), which is not unexpected given that deletions of SM cluster transporter genes have previously been shown to inhibit SM production.[22,23]

The metabolite profiles of several other double mutant strains provided important clues for the HAS biosynthetic pathway. Deletion of the cytochrome P450

(OE::*hasAΔhasH*) abolished HAS production and instead yielded large quantities of a single shunt metabolite, the prenylated Trp-Ala diketopiperazine, terezine D (**7**),[24] small amounts of which were also detected in the parent OE::*hasA* strain. In the OE::*hasA* strain, but not the P450 knock out, terezine D was accompanied by its hydroxy derivative, **8**.[24] Deletion of the putative *O*-methyltransferase (OE::*hasAΔhasC*) unexpectedly led to production of large quantities of L-7-dimethylallyltryptophan (**3**), as well as the *cis*- and *trans*- isomers of L-7-(3-methylbutadienyl)tryptophan (**4a**, **4b**), whereas production of all diketopiperazine derivatives was completely abolished. L-configuration of compounds **3**, and **4a/b** was confirmed by converting them to their corresponding Mosher amides (**Figure A4**).[25] In contrast, the NRPS deletant (OE::*hasAΔhasD*) did not accumulate prenylated tryptophan derivatives such as **3** or **4a/b**, suggesting that prenylation of tryptophan may not precede formation of NRPS-tethered tryptophan. Lastly, the FAD deletant (OE::*hasAΔhasG*) yielded the previously described *A. terreus* metabolite, astechrome (**2**) instead of HAS (**Figure 1.2b** and **Table A4**), suggesting that HasG converts the prenyl to a methylbutadienyl sidechain. Bioinformatic examination of the *A. terreus* genome revealed a gene cluster with close homology to the *A. fumigatus has* cluster; however, the *A. terreus* cluster lacks a gene encoding a homolog of the FAD HasG (**Figure A5**), supporting that HasG is required for introducing the diene moiety in the side chain (**Figure 1.2a**). We further noted that isolated samples of the *cis*-isomer **4b** slowly isomerized to **4a**, suggesting that HasG produces *cis*-7-(3-methylbutadienyl)indole derivatives, which then slowly convert to the corresponding *trans*-isomers such as **4a** and **9a** during extraction and isolation. Therefore, hexadehydroastechrome (**1**) is shown as *cis* in Figure 1.1.

Model for HAS biosynthesis: These results are consistent with a biogenesis model in which the NRPS is loaded with alanine and tryptophan, followed by prenylation of the tethered tryptophan or the resulting tethered Trp-Ala-dipeptide (**Figure 1.2a**). Deletion of the P450 *hasH* appears to stall further processing of the tethered dipeptide, which then undergoes spontaneous cyclization to form the diketopiperazine, terezine D, the only *has* pathway metabolite detected in the OE::*hasAΔhasH* strain. Our observation that the OE::*hasAΔhasC* (MT) mutant strain does not accumulate diketopiperazines, but instead yields large quantities of prenylated tryptophan derivatives, suggests a model in which HAS biosynthesis continues via tethered intermediates as shown in Figure 1.2a, as opposed to an alternative pathway involving further modification of terezine D-derived diketopiperazines. In this model, the presence of prenylated tryptophan derivatives **3** and **4a/b** could result from enzymatic hydrolysis of a tethered, partially oxidized dipeptide, for

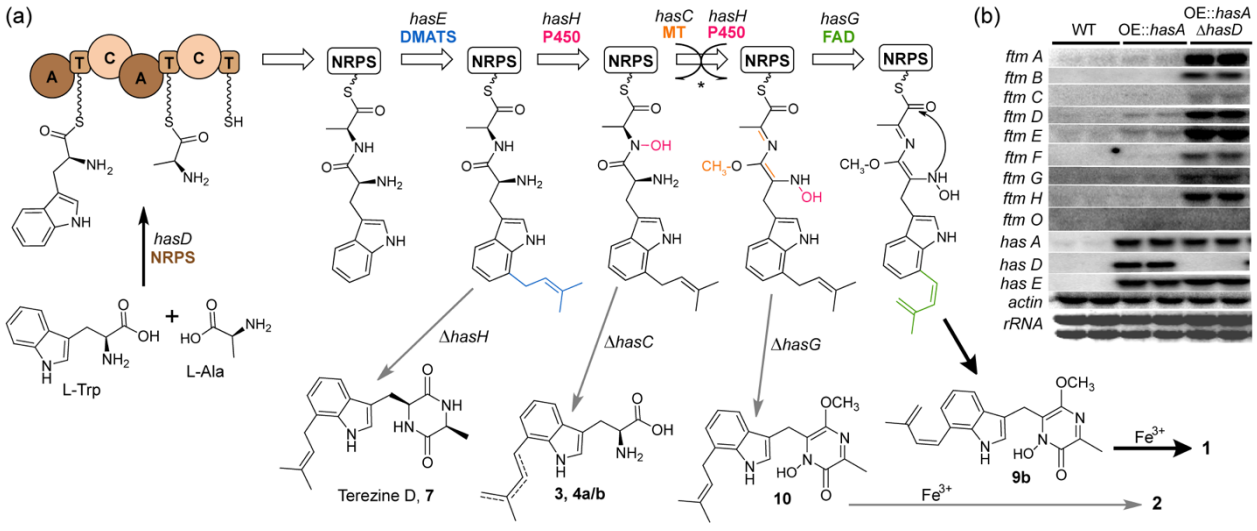


Figure 1.2: Model for HAS biogenesis and crosstalk with the *ftm* gene cluster. (a) Putative *has* biosynthetic pathway (A = adenylation, T = thiolation, C = condensation domains[16,17]), (*) order of enzyme activity unclear. (b) *ftm* gene cluster expression in two has mutants.

example by the DMATS HasE, which has previously been reported to hydrolyze linear dipeptides such as of H-L-Trp-L-Ala-OH, but not diketopiperazines.[26] Deletion of the prenyltransferase (DMATS, HasE) led to complete abolishment of HAS production and no intermediates or shunt metabolites were observed.

Surprisingly, we found that the NRPS deletant (OE::*hasAΔhasD*), although producing no HAS intermediates, yielded large amounts of metabolites derived from the partially characterized *ftm* (fumitremorgin) NRPS cluster, including brevianamide F, fumitremorgin C and demethoxy-fumitremorgin C (**5a**, **5b**, **6**, Figure 1c). Fumitremorgin biosynthesis also involves prenylation of tryptophan derivatives,[18] however, the *ftm* cluster does not contain a pathway-specific transcription factor. Examination of *ftm* gene expression in wild type (WT), OE::*hasA*, and OE::*hasAΔhasD* strains showed greatly increased *ftm* cluster gene expression in the latter strain (Figure 2b), consistent with the observation of increased *ftm*-metabolite production in OE::*hasA hasD*. These results suggest that tryptophan and prenylation activity is shunted toward fumitremorgin biosynthesis when the *has* pathway is shut down in OE::*hasA* background. Interestingly, fumitremorgin production was also increased in NRPS deletants of the fumigaclavine C pathway,[27] possibly due to a similar mechanism.

Invasive aspergillosis and the *has* biosynthetic gene cluster, a mouse model: To gain initial insight into the possibility that HAS might contribute to virulence, WT, OE::*hasA* and OE::*hasAΔhasD* strains were compared in a neutropenic murine pulmonary model.

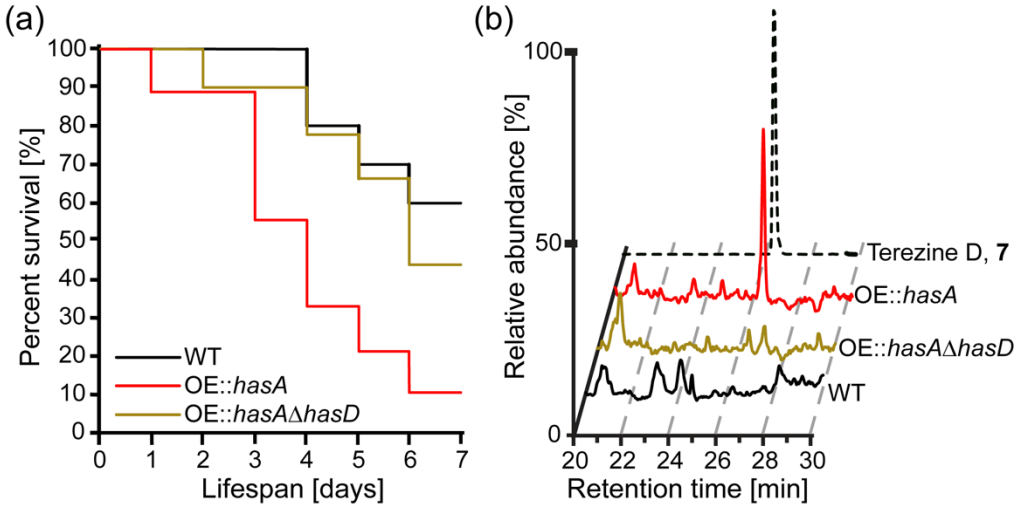


Figure 1.3: *The has* pathway affects survival in a mouse model of IA. (a) Mantel-Cox survival analysis of mice infected with WT, OE::*hasA*, or OE::*hasAΔhasD A. fumigatus* in a neutropenic pulmonary mouse model of IA. OE::*hasA* is significantly more virulent than WT or OE::*hasAΔhasD* ($P<0.05$, Log-Rank Test). There was no significant difference between WT or OE::*hasAΔhasD* strains. (b) MS ion chromatograms for Terezine D (7) standard and lung extracts from mice infected with WT, OE::*hasA*, or OE::*hasAΔhasD*.

Infection with the OE::*hasA* strain resulted in significantly higher mortality than infection with WT and OE::*hasAΔhasD* (**Figure 1.3a**). Murine lung tissue from all treatments were examined for the presence of *has* metabolites. Whereas none of the identified *has*-dependent metabolites could be detected in the WT- and OE::*hasAΔhasD*-infected samples, tissue samples from OE::*hasA*-inoculated mice revealed the presence of terezine D (**Figure 1.3b**)[24]. HAS was not detected in the OE::*hasA* lung extracts, however, based on our experience with isolating the iron(III) complex from fungal cultures, this main product of the *has* pathway is likely too unstable to permit isolation from animal tissue.

Our observation that in laboratory culture OE::*hasA*Δ*hasD* produces large amounts of fumitremorgins suggests the possibility that crosstalk with other metabolic pathways may in part be responsible for the increased virulence of OE::*hasA* in mice. However, fumitremorgins could not be detected in any of the mouse lung tissue samples, and production of gliotoxin, another metabolite contributing to *A. fumigatus* virulence,[11] was not increased in either the OE::*hasA* or OE::*hasA*Δ*hasD* samples relative to WT (**Figure A6**).

Conclusions: We show that microarray analysis coupled with comparative metabolomics using knockout (KO) and overexpression *A. fumigatus* mutants enables identification of potential virulence factors of a pathogenic fungus and proposing a tentative biosynthetic pathway. Examination of available fungal genomes indicates that *has*-related gene clusters are found only in a few opportunistic human pathogenic Aspergilli and related dermatophytic fungi of the *Trichophyton* and *Arthroderma* genera (**Figure A5**), providing additional motivation to further investigate the impact of HAS and related metabolites on virulence. Because of its iron-binding properties, we considered the possibility that **9b** could be acting as a siderophore, as siderophores are well known virulence factors in *A. fumigatus* and other fungi.[28] However, for entropic reasons it seems unlikely that the hydroxamic acid **9b** is able to compete with potent multidentate siderophores such as triacetylfusarinine C (TAFC) or ferricrocin.[28] Correspondingly, we found that neither **9b** or **1** were produced by OE::*hasA* cultures grown in iron-deficient medium, whereas TAFC production was strongly upregulated (**Figure A7**). Nevertheless, the iron requirement for HAS biosynthesis suggests that the *has* pathway and siderophore production may

interact. In addition, the highly lipophilic trimer HAS may function as a vehicle for the delivery of its much more polar monomer **9** and/or serve to inflict oxidative damage.

REFERENCES

1. Latge, J. P. Aspergillus fumigatus and aspergillosis. *Clin Microbiol Rev* **12**, 310-350, (1999).

2. Bok, J. W. *et al.* LaeA, a Regulator of Morphogenetic Fungal Virulence Factors. *Eukaryot Cell* **4**, 1574-1582, (2005).

3. Amaike, S. & Keller, N. P. Distinct Roles for veA and LaeA in Development and Pathogenesis of *Aspergillus flavus*. *Eukaryot. Cell* **8**, 1051-1060, (2009).

4. Wiemann, P. *et al.* FfVel1 and FfLae1, components of a velvet-like complex in Fusarium fijukuroi, affect differentiation, secondary metabolism and virulence. *Mol. Microbiol.* **77**, 972-994, (2010).

5. Wu, D., Oide, S., Zhang, N., Choi, M. Y., Turgeon, B. G. ChLae1 and ChVel1 Regulate T-toxin Production, Virulence, Oxidative Stress Response, and Development of the Maize Pathogen *Cochliobolus heterostrophus*. *PLoS Pathog* **8**, e1002542, (2012).

6. Perrin, R. M. *et al.* Transcriptional Regulation of Chemical diversity in Aspergillus fumigatus by LaeA. *PLoS Pathog* **3**, e50, (2007).

7. Scharf, D. H. *et al*. A dedicated glutathione S-transferase mediates carbon-sulfur bond formation in gliotoxin biosynthesis. *J Am Chem Soc* **133**, 12322-12325, (2011).

8. Forseth, R. R. *et al*. Identification of Cryptic Products of the Gliotoxin Gene Cluster Using NMR-Based Comparative Metabolomics and a Model for Gliotoxin Biosynthesis. *J Am Chem Soc* **133**, 9678-9681, (2011).

9. Scharf, D. H. *et al*. Epidithiol Formation by an Unprecedented Twin Carbon-Sulfur Lyase in the Gliotoxin Pathway. *Angew Chem Int Ed Engl* **51**, 10064-10068, (2012).

10. Nicolaou, K. C. *et al*. Synthesis and Biological Evaluation of Epidithio-, Epitetrathio-, and bis-(Methylthio)diketopiperazines: Synthetic Methodology, Enantioselective Total Synthesis of Epicoccin G, 8,8'-epi-ent-Rostratin B, Gliotoxin, Gliotoxin G, Emethallicin E, and Haematocin and Discovery of New Antiviral and Antimalarial Agents. *J Am Chem Soc* **134**, 17320-17332, (2012).

11. Dagenais, T. R., Keller, N. P. Pathogenesis of Aspergillus fumigatus in Invasive Aspergillosis. *Clin Microbiol Rev* **22**, 447-465, (2009).

12. Sugui, J. A., *et al*. Genes Differentially Expressed in Conidia and Hyphae of Aspergillus fumigatus upon Exposure to Human Neutrophils. *PLoS One* **3**, e2655, (2008).

13. Willger, S. D., *et al*. A Sterol-Regulatory Element Binding Protein Is Required for Cell Polarity, Hypoxia Adaptation, Azole Drug Resistance, and Virulence in Aspergillus fumigatus. *PLoS Pathog.* **4**, e1000200, (2008).

14. Vodisch, M., *et al*. Analysis oft he Aspergillus fumigatus Proteome Reveals Metabolic Changes and the Activation oft he Pseurotin A Biosynthesis Gene Cluster in Response to Hypoxia. *J Proteome Res* **10**, 2508-2524, (2011).

15. Cramer, R. A., Jr., Stajich, J. E., Yamanaka, Y., Dietrich, F. S., Steinbach, W. J., Perfect, J. R. Phylogenomic analysis of non-ribosomal peptide synthetases in the genus Aspergillus. *Gene* **383**, 24-32, (2006).

16. Stack, D., Neville, C., Doyle, S. Nonribosomal peptide synthesis in Aspergillus fumigatus and other fungi. *Microbiology* **153**, 1297-1306, (2007).

17. Kremer, A., Westrich, L., Li, S. M. A 7-dimethylallyltryptophan synthase from Aspergillus fumigatus: overproduction, purification and biochemical characterization. *Microbiology* **153**, 3409-3416, (2007).

18. Winter, J. M.; Behnken, S.; Hertweck, C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 22.

19. Forseth, R. R., Schroeder, F. C. NMR-spectroscopic analysis of mixtures: from structure to function. *Curr Opin Chem Biol* **15**, 38-47, (2011).

20. Pungaliya, C., *et al*. A shortcut to identifying small molecule signals that regulate behavior and development in Caenorhabditis elegans. *Proc Natl Acad Sci USA* **106**, 7708-7713, (2009).

21. Arai, K., Sato, S., Shimizu, S., Nitta, K., Yamamoto, Y. Metabolic Products of Aspergillus terreus. VII. Astechrome: an Iron-containing Metabolite of the Strain IFO 6123. *Chem Pharm Bul* **29**, 1510-1517 (1981).

22. Wiemann, P., *et al*. Biosynthesis oft he red pigment bikaverin in Fusarium fujikuroi: genes, their function and regulation. *Mol Microbiol* **72**, 931-946, (2009).

23. Lee, S., Son, H., Lee, J., Lee, Y. R., Lee, Y. W. A putative ABC transporter gene, ZRA1, is required for zearalenone production in Gibberella zeae. *Curr Genet* **57**, 343-351, (2011).

24. Wang, Y., Gloer, J. B., Scott, J. A., Malloch, D. Terezines A-D: New Amino Acid-Derived Bioactive Metabolites from the Coprophilous Fungus Sporormiella teretispora. *J Nat Prod* **58**, 93-99, (1995).

25. Dale, J. A., Mosher, H. S. Nuclear Magnetic Resonance Enantiomer Reagents. Configurational Correlations via Nuclear Magnetic Resonance Chemical Shifts of Diastereomeric Mandelate, *O*-Methylmandelate, and α-Methoxy-α-trifluoromethylphenylacetate (MTPA) Esters. *J Am Chem Soc* **95**, 512-519, (1973).

26. Kremer, A., Li, S. M. Tryptophan Aminopeptidase Activity of Several Indole Prenyltransferases From Aspergillus Fumigatus. *Chem Biol* **15**, 729-738, (2008).

27. O'Hanlon, K. A., *et al*. Nonribosomal peptide synthetase genes pesL and pes1 are essential for Fumigaclavine C production in Aspergillus fumigatus. *Appl Environ Microbiol* **78**, 3166-3176, (2012).

28. Haas, H. Iron – A Key Nexus in the Virulence of Aspergillus fumigatus. *Front Microbio* **3**, 28, (2012).

CHAPTER 2

SECOND CONDENSATION-LIKE DOMAIN IN GLIP IS ESSENTIAL FOR DIKETOPIPERAZINE CYCLIZATION IN GLIOTOXIN BIOSYNTHESIS

Abstract: One of the most well-studied metabolites and virulence factors from the human pathogen *Aspergillus fumigatus* is gliotoxin, an epidithiodiketopiperazine (ETP) toxin. Despite extensive research on gliotoxin's biosynthesis, bioactivity and ecology, the enzymatic activity responsible for diketopiperazine (DKP) formation en route to gliotoxin has remained uncharacterized. Here we show that the second condensation (C) domain in the non-ribosomal peptide synthetase (NRPS) GliP functions as a cyclization (C_T) domain, which is essential for DKP formation and ultimately gliotoxin biosynthesis. *In vivo* truncation of GliP to remove the C_T domain resulted in a complete loss of gliotoxin production, as well as all intermediates and shunt metabolites in the *gli* pathway. *In vitro* reconstitution of the C_T domain confirmed its essential role in cyclization of dipeptidyl substrates. We further show that a single point mutation of a conserved catalytic histidine eliminated cyclization activity of the C_T domain both *in vitro* and *in vivo*. These results indicate gliotoxin biosynthesis proceeds via tailoring modifications that occur on a linear dipeptide scaffold attached to GliP, rather than on the released cyclo(L-Phe-L-Ser). The GliP C_T domain belongs to the same class of cyclization domains involved in fumiquinazoline biosynthesis, however had not previously been annotated as such. Lastly, this work highlights a general strategy for the biosynthesis of modified DKPs derived from NRPSs with similar domain architecture to GliP.

Introduction: Gliotoxin was discovered in 1936 by Weindling and Emerson and since then has been the subject of several hundred publications[1,2]. After some debate the correct structure was elucidated in 1958 by Johnson and Woodward, which was later confirmed by X-ray crystallography (**1**, **Figure 2.1b**)[3,4]. Throughout the later part of the 20th century many total syntheses of **1** were undertaken, as well as isolations from several fungal species and widespread biological activity testing[5-8]. Suspected of playing a key role in *Aspergillus fumigatus'* pathogenicity towards humans and livestock, understanding the biogenesis of gliotoxin became an important goal for natural product chemists and fungal biologists alike[9,10]. Immediately after *A. fumigatus* was first sequenced in 2005, Gardiner and Howlett discovered the genetic basis of **1** by using bioinformatics analyses of gene expression data to correctly predict its biosynthetic gene cluster (BGC), the *gli* cluster[11]. Because of the availability of sequencing data and the human health implications, nearly all of the work on the *gli* cluster has since been conducted in *A. fumigatus*[12].

The core enzyme in the *gli* cluster, GliP, is a nonribosomal peptide synthetase (NRPS) comprised of seven domains, namely two adenylation (A) domains, two condensation (C) domains, and three thiolation (T) domains, previously annotated and ordered as A_1-T_1-C_1-A_2-T_2-C_2-T_3 (**Figure 2.1a**)[13]. The early steps in gliotoxin formation are as follows: A_1 activates phenylalanine and A_2 activates serine, which are tethered to T_1 and T_2, respectively, then condensed by the C_1 domain, forming T_2 attached L-Phe-L-Ser dipeptide (**Figure 2.1a**). While the mechanism and role of A_1-T_1-C_1-A_2-T_2 domains in the formation of T_2-tethered L-Phe-L-Ser is straightforward, the ensuing modification steps and release from GliP are not well defined and cannot be predicted simply from the

annotation of GliP and *gli* cluster tailoring genes. In terms of GliP, the C_2 and T_3 domains in particular have no obvious function, since an additional amino acid is not present in the structure of gliotoxin and the absence of a third A domain would mean involvement of a stand-alone A domain, for which there are no candidate genes in or around the *gli* cluster. In an attempt to clarify the biosynthetic function of each domain in GliP, Balibar and Walsh recombinantly expressed GliP in *Escherichia coli* in 2006[13]. They confirmed the canonical function of A_1-C_1-T_1-A_2-T_2 towards the formation of L-Phe-L-Ser through the observation of cyclo(L-Phe-L-Ser) (**2**), which they hypothesized resulted from the spontaneous cyclization of L-Phe-L-Ser from the T_2 domain (**Figure 2.1c**). They were unable to discern the function of C_2 and T_3, as mutational inactivation of these domains did not impact the rate of formation of **2** and no new L-Phe-L-Ser related products were observed. Shortly after the *in vitro* reconstitution of GliP several efforts were undertaken by this lab and others to describe the full *gli* biosynthetic pathway using both *in vivo* and *in vitro* approaches.

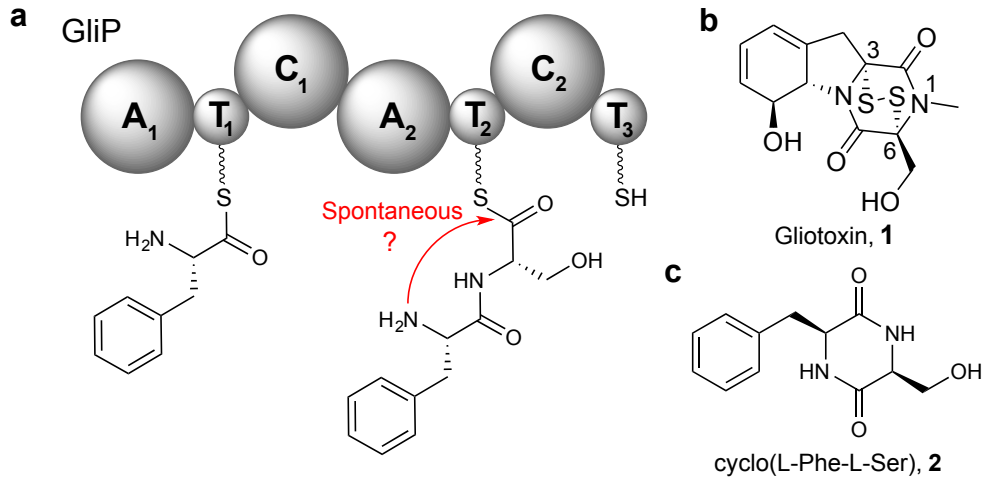


Figure 2.1. Domain architecture of GliP and DKP formation in question. (**a**) GliP loads L-Phe on T_1 and L-Ser on T_2, which are condensed together by the C_1 domain to form T_2-tethered L-Phe-L-Ser (A = adenylation, T = thiolation, C = condensation domain) (**b**) Structure of the final product of the *gli* pathway, Gliotoxin, **1**. (**b**) Structure of early released *gli* pathway metabolite cyclo(L-Phe-L-Ser), **2.**

In 2011 our laboratory provided the most complete picture of gliotoxin biosynthesis to date[14]. Using DANS-based comparative metabolomics of WT *A. fumigatus* versus knock-out of *gliZ* (the *gli* cluster transcriptional regulator), six previously unidentified gliotoxin intermediates and shunt metabolites were characterized (red structures, **Figure B1**). These novel *gli* pathway metabolites were consistent with a biosynthetic model where tailoring modifications occur to a GliP-tethered linear L-Phe-L-Ser dipeptide (**Figure B1**). Our proposal for a tethered model of gliotoxin biosynthesis was in contradiction with prior proposals, where the released product of GliP was speculated to be an unmodified **2**, rather than a modified DKP that had already been elaborated by *gli* cluster tailoring enzymes (**Figure B1**)[15]. Evidence for the untethered model (where **2** is

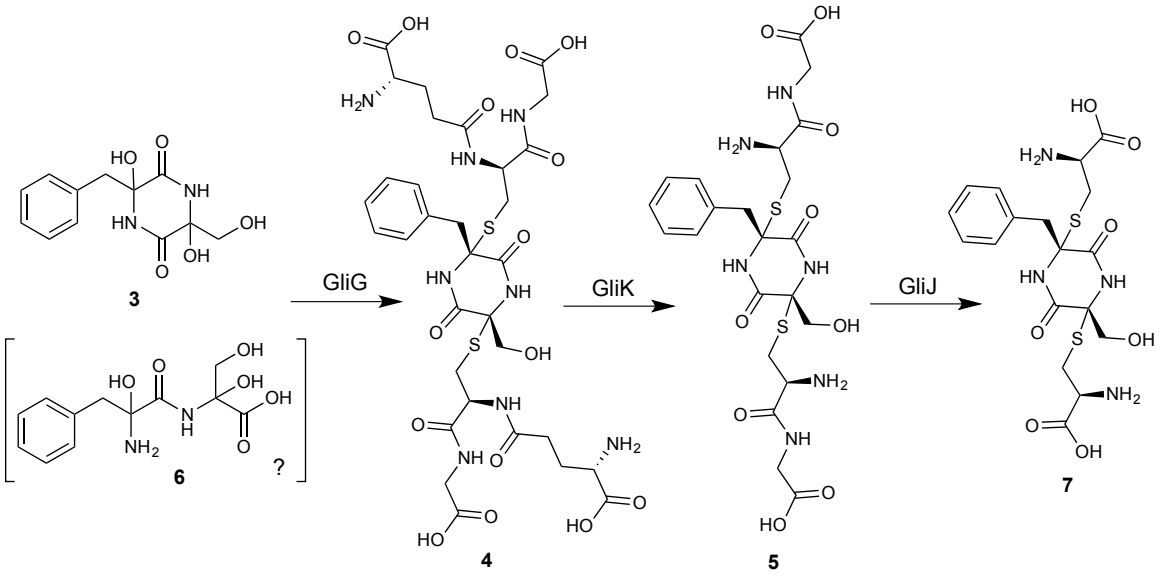


Figure 2.2. In vitro characterization of gli tailoring enzymes by the Hertweck group. GliG converts **3** to **4** by addition of two GSH groups at each alpha carbon. GliK cleaves Glu from GSH to convert **4** to **5**, then GliJ cleaves Gly resulting in **7**.

the true product of GliP) is based in part on the original *in vitro* characterization of GliP, which doesn't account for the possibility of altered GliP activity when *gli* cluster tailoring enzymes are present and physically interacting, a situation that very likely occurs *in vivo*.

Additional support for the untethered model came shortly after our lab published on the full *gli* pathway in 2011 when Hertweck and coworkers published the characterization of GliG, a Glutathione *S*-Transferase in the *gli* cluster[16].

Deletion of GliG resulted in the accumulation of **3** and complete abrogation of production of **1** in *A. fumigatus* (**Figure 2.2**). Extracts from an *A. fumigatus* ΔGliG mutant that were treated with recombinant GliG resulted in the production of glutathione (GSH) derivative **4**, which derived from intermediate **3** (**Figure 2.2**). This was taken as evidence that the substrate for both GliC, a cytochrome p450 hydroxylase (responsible for the oxidation of **2** to **3**), and GliG was indeed **2**, however, **3** could also have resulted from an oxidation of GliP-tethered L-Phe-L-Ser (**Figure B1**). If this was the case then it would be inconsistent with the next enzymatic step where GliK converts **4** to **5**, however, Hertweck and coworkers did not test substrate specificity to show GliG preferred the cyclic substrate **4** over a linear dipeptide, such as **6** (**Figure 2.2**). Hertweck and coworkers later described the sulfurization steps involved in gliotoxin biosynthesis in 2013, which were also contrary

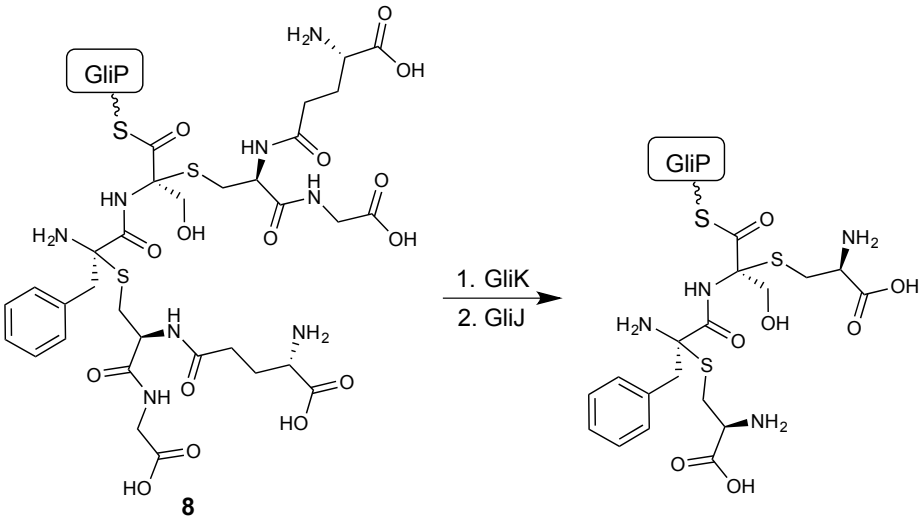


Figure 2.3. Alternative hypothetical enzymatic activity for GliK and GliJ peptidases. GliK and GliJ dipeptidases convert GliP-tethered L-Phe-L-Ser-diGSH to GliP-tethered L-Phe-L-Ser-diCys.

to our original model for **1** biogenesis[17]. They showed that GliK and GliJ converted **4** to cyclo(L-Phe-L-Ser)-Cys **7**, (**Figure 2.2**). Their characterization of the sulfurization steps did not account for the oxidation steps prior to sulfurization, as well as substrate specificity of the enzymes required for sulfurization. Considering GliJ and GliK are peptidases, which are known to exhibit high-levels of substrate promiscuity[18], it seems possible that they could also function in accordance with the tethered model for gliotoxin biosynthesis that we initially reported. However, the creation of a GliP-tethered intermediate, such as **8** to test this hypothesis is extremely difficult (**Figure 2.3**).

Perhaps, the most unsatisfying aspect of the untethered model for **1** biosynthesis is the neglect of the C_2 and T_3 domains in GliP. If they were truly unnecessary, then it seems unlikely they would have been maintained through evolution. The C_2 and T_3 domains are encoded for by large genomic sequences and **1** is important for fungal fitness – if they were unnecessary, eventually a successful species would arise where they were missing. However, this is not the case, as GliP is highly conserved across many fungal species, which points to an important and unrealized function of the C_2 and T_3 domains. An indication for the possible function of the GliP C_2 and T_3 domains came in 2012 when Tang and Walsh discovered that some terminal condensation domains were actually cyclization (C_T) domains[19]. These fungal specific C_T domains were shown to be necessary for the macrocyclization of linear peptidyl precursors produced by a variety of fungal NRPSs. They showed the activity of the C_T domain to be dependent on the conserved residues (SHXXXDXXS/T), and specifically the conserved histidine. Tang and Walsh re- annotated many NRPSs involved in the biosynthesis of macrocyclic peptides as having essential terminal C_T domains, rather than simply C domains. They did not

however, re-annotate the C_2 domain of GliP, despite the fact that it too contains the conserved sequence of the C_T domain (**Figure B2**)[19]. This is likely because the C_2 is not a terminal domain in GliP, whereas, Tang and Walsh specifically characterized a terminal C_T domain from the trimodular NRPS TqaA[20]. The homology of C_2 in GliP to the published C_T domain strongly suggested it serves as an essential cyclization domain toward **1** biosynthesis. Additionally, the results described in Chapter 1 of this dissertation provided more support for C_2 actually being C_T and **1** biosynthesis proceeding via a tethered model.

In Chapter 1 the characterization of the *has* BGC is described. When we initially set out to understand the *has* pathway, HasD, the *has* cluster NRPS, was annotated in NCBI as another GliP[18]. This was because HasD and GliP share the same domain architecture of A_1-T_1-C_1-A_2-T_2-C_2-T_3. Despite sharing high homology for their NRPSs, the *has* and *gli* clusters make completely different small molecule products, with the exception of both products containing the structural feature of a modified DKP. In the *has* biosynthetic pathway deletion of HasC, the *has* O-methyltransferase tailoring enzyme, leads to the accumulation of compounds **3-4a/b** from Chapter 1. We hypothesized **3-4a/b** from Chapter 1 resulted from a stalled biosynthesis of a tethered dipeptide attached to HasD, which was cleaved by the *has* cluster encoded dipeptidase, HasE, as described in Chapter 1. Fortunately, another lab conducted the *in vitro* characterization of HasE and showed it does indeed cleave linear dipeptides and importantly, not cyclic DKPs, evidence in favor of the tethered model for the *has* pathway[21]. Analogously, the *gli* cluster also contains a dipeptidase, GliJ, which was previously shown to cleave glycine from **5**, in the second step of four sulfurization steps in the *gli* pathway (**Figure 2.2**). We suspected that in addition to cleaving **5**, GliJ activity would also serve a function

analogous to that of HasE and cleave linear L-Phe-L-Ser in a similar recycling-type mechanism as observed for HasE in the *has* pathway.

Here we show that the C_2 domain in GliP is actually a C_T domain that is essential for gliotoxin (**1**) biosynthesis, as predicted based on the homology to other C_T domains. Construction of an *A. fumigatus* point mutant for the conserved catalytic histidine (GliP-H1754A) predicted to be essential for cyclization activity resulted in complete loss of gliotoxin production and all *gli*-pathway metabolites. *In vitro* reconstitution of a truncated GliP-ΔC_2T_3 protein resulted in a dramatic decrease in rate of formation of **2**. Our results show that the GliP C_2 domain is necessary for cyclizing an unmodified L-Phe-L-Ser substrate and that a modified GliP-T_3-tethered L-Phe-L-Ser likely represents the true substrate, which is what occurs *in vivo*. Further supporting the cyclization role of the C_2 domain we also show that GliJ's substrate specificity is not limited to Cys-Gly dipeptides and that it may serve an additional recovery-type role by cleaving a stalled L-Phe-L-Ser that is still attached to GliP.

***In vivo* truncation of GliP abolished gliotoxin production:** In order to determine the biosynthetic function of the GliP C_2 and T_3 domains *in vivo* we first constructed truncation mutants in *A. fumigatus* for each of these domains separately, GliP-ΔT_3 and GliP-ΔC_2T_3 (**Appendix B**). To assess the impact of each truncation mutant on gliotoxin biosynthesis we subjected whole metabolome extracts to a liquid chromatography-high resolution mass spectrometry (LC-HRMS) comparative metabolomics analysis and surveyed for gliotoxin and two other abundant *gli* pathway metabolites (**1-2**, **9-10**, **Figure 2.4 and Figure B1**). We were specifically interested in **2**, as production of this compound should

not have been hindered by elimination of the terminal C_2 and T_3 domains if the formation was spontaneous from T_2, as previously reported. LC-HRMS comparison of whole metabolome extracts from WT (Af293), GliP-ΔT_3, and GliP-ΔC_2T_3 revealed a complete loss of gliotoxin and pathway intermediates (**1-2**, **9-10**) in the GliP-ΔC_2T_3 strain (**Figure 2.4a-d**). In extracts from the GliP-ΔT_3 mutant strain we were unable to detect **1** and the most abundant *gli* pathway metabolites **2** and **9**, however we observed trace quantities of **10** (99.3 % less than Af293, n=4, p=0.002), the detoxification product of **1** (**Figure 2.4a-d**). These data show that biosynthesis of **1** proceeds without the T_3 domain of GliP, however, only at less than one percent production compared to WT (Af293).

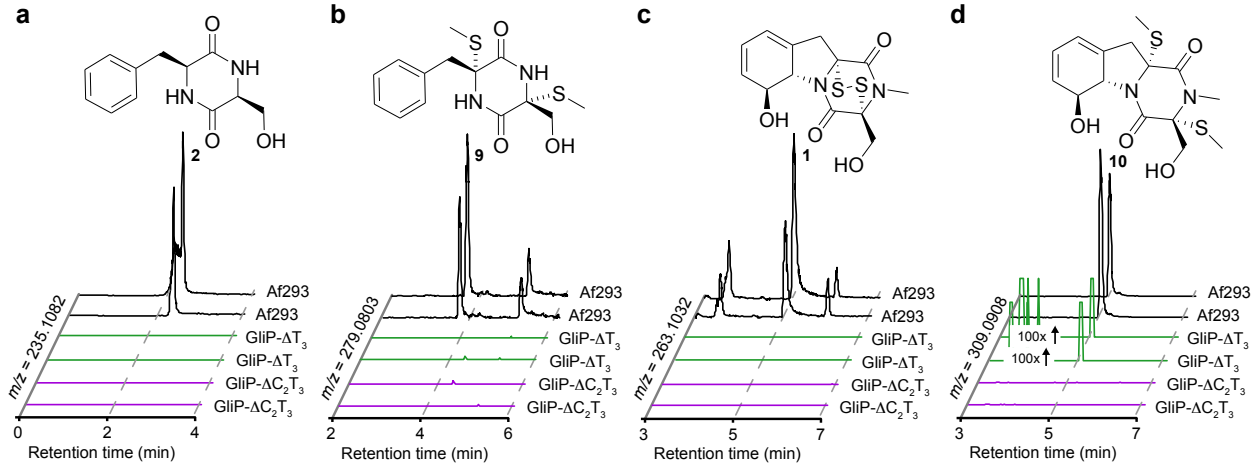


Figure 2.4. LC-HRMS analysis of whole metabolome extracts from WT and mutant *A. fumigatus* strains. (**a-d**) Representative overlaid extracted-ion-chromatograms for **1**, **2**, **9** and **10** in Af293 (black lines), GliP-ΔT_3(green lines), and GliP-ΔC_2T_3 (purple lines).

Gliotoxin (**1**) serves as a positive feedback loop for its own production through regulation of *gli* cluster expression[22]. The inability of the GliP-ΔC_2T_3 mutant strain to produce **1** resulted in a complete loss of *gli* cluster expression (**Figure B3**). Similarly, the reduction of gliotoxin (**1**) production observed for the GliP-ΔT_3 mutant strain resulted in decreased *gli* cluster expression but not a complete abolishment as was seen for the GliP-ΔC_2T_3 mutant strain (**Figure B3**). Our initial analysis of the whole metabolome

extracts from the GliP-ΔC_2T_3 and GliP-ΔT_3 mutant strains was therefore biased, since the lack of *gli* cluster expression could account for the observed metabolite profiles. We rescued *gli* cluster expression by growing cultures supplemented with exogenous **1** (20 μg/mL in GMM) (**Figure B4**). We did not do a quantitative comparison to existing data[22] for *gli* cluster rescue, rather we qualitatively determined by northern blot that the feeding of **1** had brought back *gli* cluster expression to similar levels as observed in the WT strain,

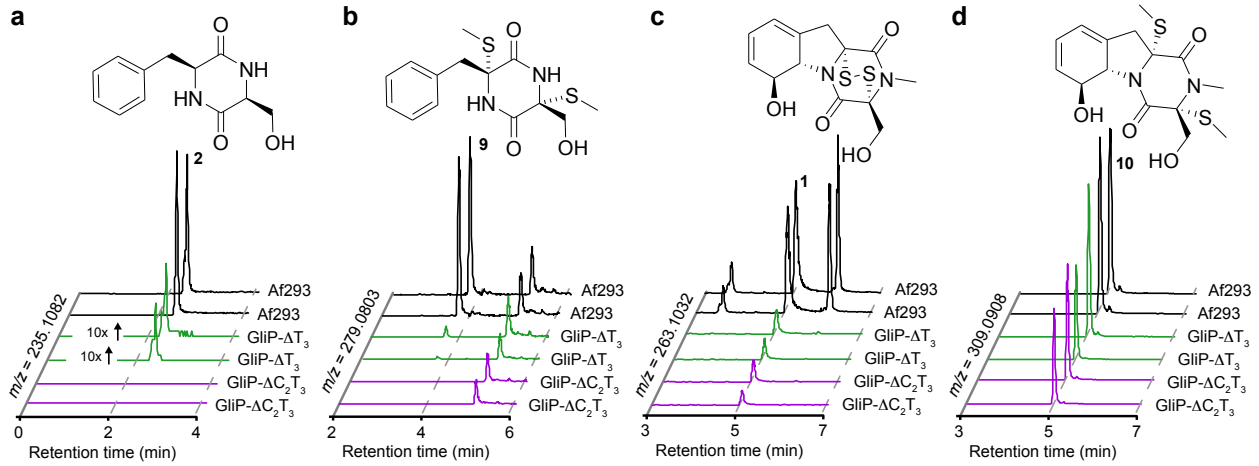


Figure 2.5. LC-HRMS analysis of whole metabolome extracts from WT and mutant *A. fumigatus* strains supplemented with exogenous 1 during growth. (**a-d**) Overlaid extracted-ion-chromatograms for **1**, **2**, **9** and **10** in Af293 (black lines), GliP-ΔT_3(green lines), and GliP-ΔC_2T_3 (purple lines)

which should enable observation of pathway metabolites. Additionally, we did not explore effects of gliotoxin feeding beyond the direct impact on *gli* cluster expression, as such an exploration would be outside the scope of this work. LC-HRMS comparison of cultures supplemented with **1** showed that despite rescued expression, GliP-ΔC_2T_3 did not produce any *gli* pathway metabolites (**Figure 2.5a-d**). The GliP-ΔT_3 strain however, produced significantly less **2** (95.7 % less than Af293, n = 3, $p < 0.0001$) and the tailored metabolite, **9** (96.3 % less than Af293, n = 4, $p = 0.026$) (**Figure 2.5a-b**). These results suggest that in the absence of T_3, low-level production (3-5 % of WT levels) of **1** and other *gli* pathway metabolites is possible via tailoring steps on a T_2-tethered intermediate, which

is cyclized by C_2, albeit with extremely reduced activity. This further supports the hypothesis that T_3 is a bona fide third peptidyl carrier domain and serves as the primary location for linear substrates to be modified by *gli* cluster tailoring enzymes, rather than a remnant domain with no biosynthetic function.

Adenylation activity of GliP-ΔC_2T_3 and GliP-ΔT_3 A_1 and A_2 domains is equal or greater than GliP-WT: One concern regarding the loss of *gli* pathway metabolites in the GliP-ΔC_2T_3 and GliP-ΔT_3 strains is that by truncating GliP we changed the overall structure significantly enough to impede A domain adenylation activity. This would result in L-Phe and L-Ser never being loaded onto T_1 and T_2, accounting for the observation of no or minimal amounts of *gli* pathway metabolites. To address this concern we heterologously expressed and purified GliP-WT, GliP-ΔC_2T_3 and GliP-ΔT_3 from *E. coli* (**Figure B5** and **Appendix B**). We evaluated the adenylation activity of A_1 and A_2 for L-Phe and L-Ser, respectively, in all three proteins using the standard ATP-[^{32}P]pyrophosphate exchange assay[23]. The truncated proteins, GliP-ΔC_2T_3 and GliP-ΔT_3, showed adenylation activity for L-Phe 35-45 fold and 10-20 fold higher than water, respectively, whereas, GliP-WT showed 320-380 fold higher adenylation activity than water (**Figure B6** and **Appendix B**). The decrease in adenylation activity as compared to water for the truncation mutants could be caused by higher levels of ATP-hydrolysis in the shorter proteins. Comparison of L-Phe to L-His adenylation activity of GliP-WT, GliP-ΔC_2T_3, and GliP-ΔT_3 showed more comparable turn-over rates (GliP-WT = 130-170 fold increase for L-Phe, GliP-ΔT_3 = 50-80 fold increase for L-Phe, and GliP-ΔC_2T_3 = 110-140 fold increase for L-Phe, **Figure B6**). In agreement with the initial characterization of GliP by Balibar and Walsh we also

observed a roughly 5-fold higher turn-over rate for L-Phe compared to L-Ser, which was independent of the version of GliP (**Figure B6**). These results suggest that A domain activity is reduced but not eliminated by changes in the terminus of GliP.

***In vitro* product formation assays with GliP-ΔC_2T_3, GliP-ΔT_3, and GliP-WT are consistent with *in vivo* results:** Next we sought to characterize the product formation of each truncated protein by incubating the purified proteins with L-Phe and L-Ser (see **Appendix B** for experimental details). We monitored each reaction by LC-MS and measured the rate of formation of cyclo(L-Phe-L-Ser) (**2**), linear L-Phe-L-Ser (**12**), and looked for any new products that had not been found in prior analyses of GliP's *in vitro* activity. Our results for GliP-WT were consistent with the published findings, which was a slow formation of primarily **2** with small quantities of hydrolyzed linear **12** (**Figure 2.6a**). The GliP-ΔT_3 protein had slightly reduced activity as measured by the formation of **2** + **12,** (93% of GliP-WT, **Figure 2.6b and Figure B7**). Conversely, GliP-ΔC_2T_3 had dramatically reduced activity (33% of GliP-WT, **Figure B7**) and the ratio of both hydrolyzed **12** and L-Phe captured **13** to cyclic **2** was increased by 2% (**Figure 2.6c**). In addition to the previously detected products of GliP (**2** and **12**) we also detected linear L-Phe-L-Phe and L-Phe-Tris buffer adducts (**13** and **14**, respectively, **Figure 2.6b-d**). The formation of **12-14** in each assay results from insufficient offloading of T_2 or T_3 bound substrates. This effect is more pronounced with GliP-ΔC_2T_3 (20-30% more **14** observed, **Figure 2.6c**) because in the absence of the C_2 domain, the T_2 attached **12** is stalled, which increases the amount of time for T_1 attached L-Phe to be captured by another L-Phe (to form **13**), water (to form **12**), or Tris-buffer (to form **14**).

In order to determine if the cyclization activity of the C_2 domain absent in GliP-ΔC_2T_3 could be rescued by addition of free C_2T_3 we expressed and purified standalone GliP-C_2T_3 from *E. coli* (**Figure B8** and **Appendix B**). Addition of the standalone C_2T_3 showed a minor rescue (4% increase from GliP-ΔC_2T_3 by itself, **Figure B7**) in activity and reversed the ratio of linear **2** to **12** to similar levels as seen in GliP-WT, suggesting addition of free C_2T_3 decreased the amount of spontaneous hydrolysis of T_2 tethered substrate (**Figure 2.6d**). The four percent rescue in activity was not ideal to show C_2 as a standalone cyclization domain. However, a couple of inherent problems with the experiment likely impeded full activity rescue, such as, (1) The N-terminal H_6-tag on the purified standalone C_2T_3 could act as a physical barrier preventing it from interacting with

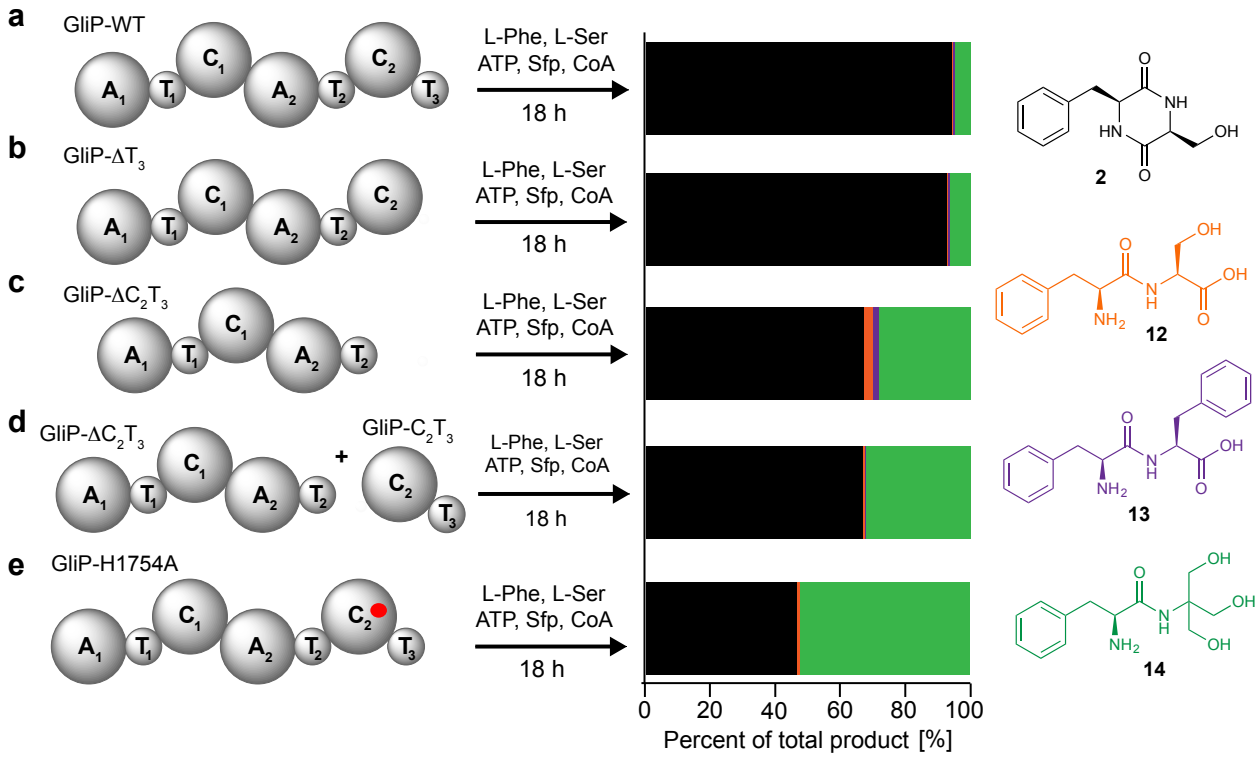


Figure 2.6. Product distribution for *in vitro* assays with GliP-WT, GliP-ΔT_3 and GliP-ΔC_2T_3. Percentage for compounds produced in each assay as **2** (black bar), **12** (orange bar), **13** (purple bar) and **14** (green bar) for (**a**) GliP-WT, (**b**) GliP-ΔT_3, (**c**) GliP-ΔC_2T_3, (**d**) GliP-ΔC_2T_3 + C_2T_3, and (**e**) GliP-H1754A.

GliP-ΔC_2T_3 in the correct orientation (we were unable to obtain soluble protein with C-terminal H_6-tag), causing a decrease in reaction rate; (2) the cyclization activity of C_2 likely prefers substrates bound to T_3 and the trans-thiolation of T_2-tethered **12** to T_3-tethered **12** would be reduced with two separate proteins. In order to determine if the substrate is indeed passed to T_3 via a trans-thiolation, we tryptic digested the product formation assay of GliP-WT and ran LC-HRMS/MS. We show that T_3 is modified with a phosphopantetheinyl arm using the phosphopantetheinyl-ejection assay, but we could not capture a T_3-tethered **12** (**Figure B9**).

Our hypothesis is that T_3 has been maintained in GliP and related NRPSs because it's a place for the linear dipeptide to be received by tailoring enzymes, which provides both space and time for elongation of newly forming **12** on T_2. This also means that the cyclization activity of C_2 should prefer a T_3 bound substrate. In order to clarify the role of T_3 *in vitro*, we heterologously expressed and purified another GliP mutant, where the first and second thiolation domains were inactivated by substitution of both serine 555 and 1582 to alanines, respectively (GliP-ΔT_1T_2). Then, we synthesized *N*-acetylcystamine (SNAC)-L-Phe-L-Ser (**15**) to enable a trans-thiolation reaction to create the T_3-tethered **12** (**Figure B10a** and **Appendix B**). It is also possible that **15** serves as a substrate for the C_2 domain, without first being attached to T_3 (**Figure B10b**). To distinguish between T_3 bound substrate (**15** post trans-thiolation) and C_2 acting on **15** itself, we conducted a control experiment without Sfp, meaning T_3 would have no posphopantetheinyl arm to

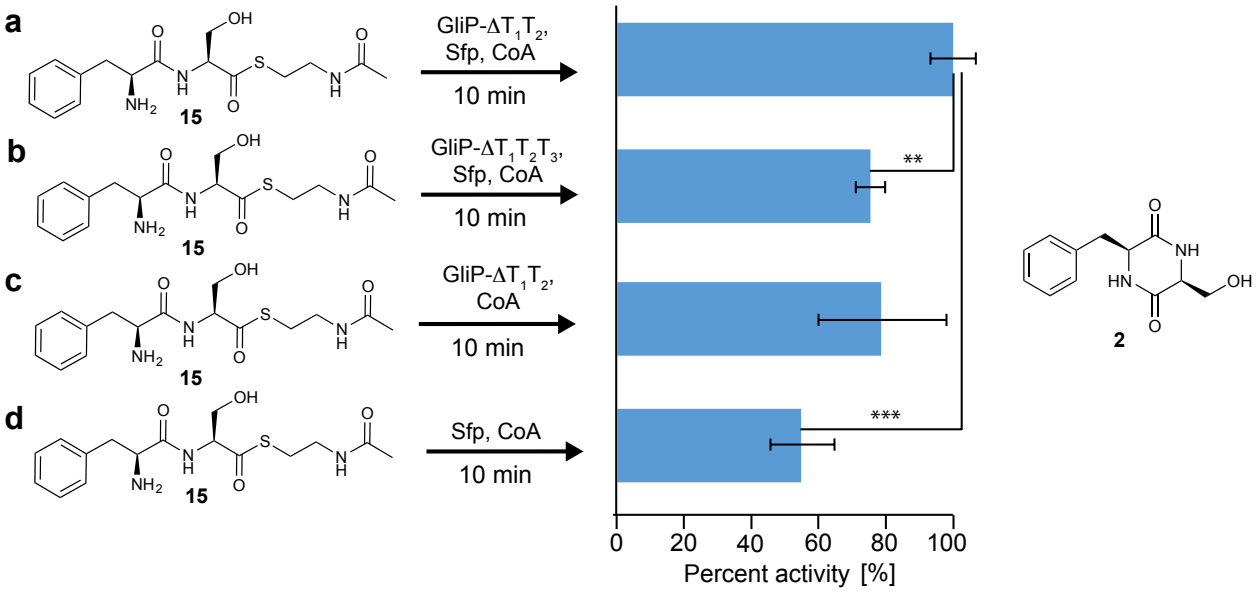


Figure 2.7. Percent cyclization of 15 to form 2. (**a**) assay with GliP-ΔT_1T_2, Sfp, CoA (**b**) assay with GliP-$\Delta T_1T_2T_3$ (**c**) control assay without Sfp, (**d**) control assay with no GliP-ΔT_1T_2. ** = p<0.01, *** = p<0.001.

capture **15**. Incubation of **15** with GliP-ΔT_1T_2, with and without Sfp, shows an increase of cyclization activity of 22 % more **2** when Sfp is present (**Figure 2.7a** compared to **Figure 2.7c**). Since, **15** spontaneously cyclizes to form **2** in solution, the background cyclization rate is high, however, the presence of T_3 significantly increases the amount of cyclic product (25% increase with T_3 present, **Figure 2.7a** compared to **Figure 2.7b**).

In the original publication by Tang and Walsh on the fungal C_T domain they described a conserved sequence that was used to re-annotate C domains to C_T domains across many fungal NRPSs. They noted that within the conserved sequence the cyclization activity could be pinned down to a single catalytic histidine. In the C_2 domain of GliP that conserved catalytic histidine is H1754 (**Figure B2**)[19]. Before taking the time to create the GliP-H1754A point mutant in *A. fumigatus* we quickly took the vector for GliP-WT and used site-directed mutagenesis to create the construct for recombinant GliP-H1754A point mutant. Expression and purification of the H_6-tagged GliP-H1754A point

mutant was carried out using previously described methods (**Appendix B**). We tested the ability of GliP-H1754A to form **12** or **2** using the same assay conditions described above. The GliP-H1754A point mutant had nearly identical product profiles to that of the GliP-ΔC_2T_3 (**Figure 2.6e**), showing that in GliP the predicted conserved catalytic histidine is essential for C_2 cyclization activity. Additionally, this provided us with the precedent to create the GliP-H1754A point mutant in *A. fumigatus* and evaluate the effect of this single point mutation on **1** biogenesis *in vivo*.

Point mutation of histidine 1754 in the C_2 domain of GliP abolishes gliotoxin production: In order to create the GliP-H1754A point mutant strain of *A. fumigatus* we first created a new ΔGliP strain, then via homologous recombination added back in GliP-WT and GliP-H1754A at the KU locus (see **Appendix B** for details), resulting in three new strains, ΔGliP, +GliP-WT and +GliP-H1754A. We grew each of these three strains in parallel in liquid media at 30 °C for five days and processed as previously described (see

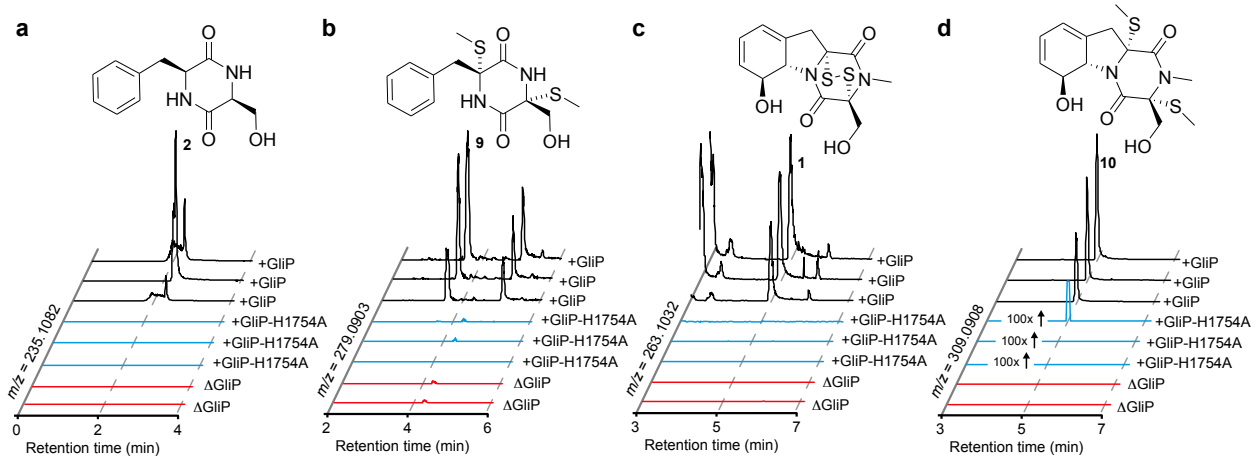


Figure 2.7. LC-HRMS analysis of whole metabolome extracts from WT and mutant *A. fumigatus* strains. (**a-d**) Overlaid extracted-ion-chromatograms for **1**, **2**, **9** and **10** in +GliP (black lines), +GliP-H1754A (blue lines), and ΔGliP (red lines).

Appendix B for details). Following whole metabolome extraction, we ran LC-HRMS and surveyed for the most abundant gliotoxin pathway metabolites (**1-2**, **9-10**). ΔGliP and +GliP-H1754A resulted in a complete loss of all *gli* pathway metabolites, whereas, the +GliP-WT, despite being in ΔGliP background, showed a normal distribution of **1-2** and **9-10** (**Figure 2.7a-d**), proving that a GliP located at the KU locus is still functional with the native *gli* cluster.

The lack of gliotoxin production in the +GliP-H1754A point mutant strain again resulted in starkly decreased expression of *gliP* and *gli* cluster tailoring genes. That the loss of a single histidine would result in complete abolishment of the gliotoxin pathway was rather surprising, since if the substrate for most *gli* cluster tailoring enzymes is in fact **2** and derivatives thereof (as has previously been described by others[12,15,24]), the loss of the histidine 1754 should not completely eliminate gliotoxin production. This alternative untethered model for gliotoxin biosynthesis in which **2** is released by spontaneous cyclization from T_2 would not be affected by the elimination of a single histidine residue in C_2. Even in the case of a tethered model for gliotoxin biosynthesis, as we have

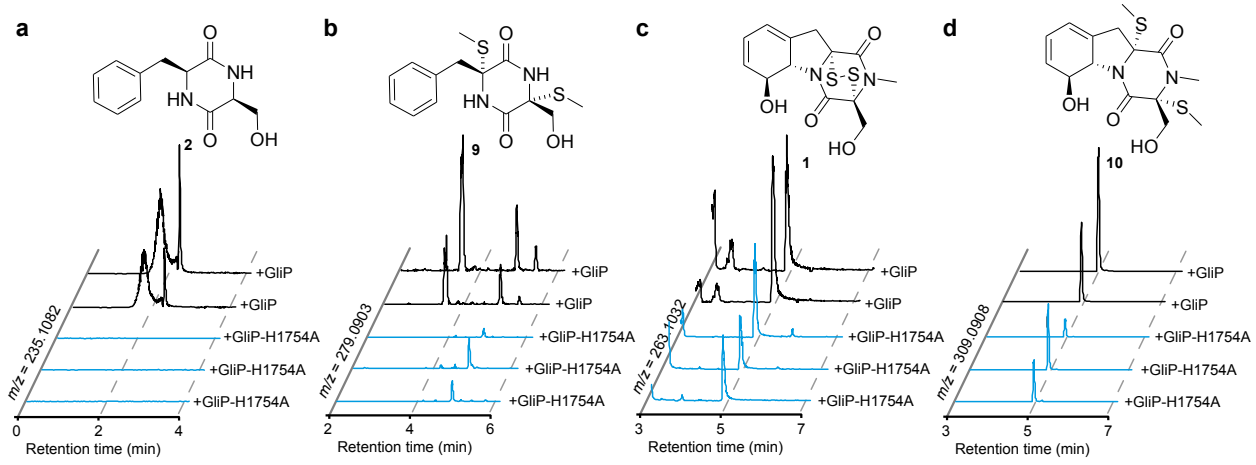


Figure 2.8. LC-HRMS analysis of whole metabolome extracts from WT and mutant *A. fumigatus* strains supplemented with exogenous 1 during growth. (**a-d**) Overlaid extracted-ion-chromatograms for **1**, **2**, **9** and **10** in +GliP (black lines), and GliP-H1754A (blue lines).

hypothesized, we expected to still observe **2**, as this intermediate is produced in large quantities by the WT enzyme (several mg/L in liquid culture, **Figure 2.4** and **Figure 2.5**). Upon feeding of exogenous **1** to +GliP-H1754A cultures to rescue *gli* cluster expression, we still observed a complete loss of *gli* pathway metabolites (**Figure 2.8a-d**). This combined with our observations from the GliP-ΔC_2T_3 and GliP-ΔT_3 strains strongly support a model for gliotoxin biosynthesis in which many of the modification steps occur while **12** is tethered to the T_3 domain.

***A. fumigatus* ΔGliP strain fed with 2 does not result in more advanced *gli* pathway metabolites:** In our publication on the *gli* pathway in 2011 we conducted a feeding experiment where **2** was added to liquid cultures of the *A. fumigatus* ΔGliP strain[25]. At the time we were unaware that in the absence of gliotoxin, the *gli* cluster tailoring enzymes were not expressed. It was, however, shown that in the ΔGliP strain, exogenous gliotoxin added to the cultures rescued *gli* cluster expression to WT levels, as described above[22]. This means in our initial feeding experiment the observation that **2** was not processed by *gli* cluster tailoring enzymes was uninformative, since they weren't expressed. To address this void in evidence we repeated the feeding experiment, this time however, in addition to **2** we also added exogenous **1**, as described above (**Appendix B**). Additionally, instead of feeding the full knock-out GliP strain (ΔGliP) where no core enzyme is present (disrupting potential physical interactions between cluster tailoring enyzmes) we instead fed the GliP-H1754A point mutant strain. Even with WT-levels of *gli* cluster expression (as evidenced by the conversion of **1** into **10**, **Figure 2.9c-d**), and all physical components of the *gli* cluster present (non-functional core enzyme available for essential non-

enzymatic interactions with tailoring enzymes), the supplemented **2** was not processed to more advanced intermediates in the *gli* pathway (**Figure 2.9a-b**).

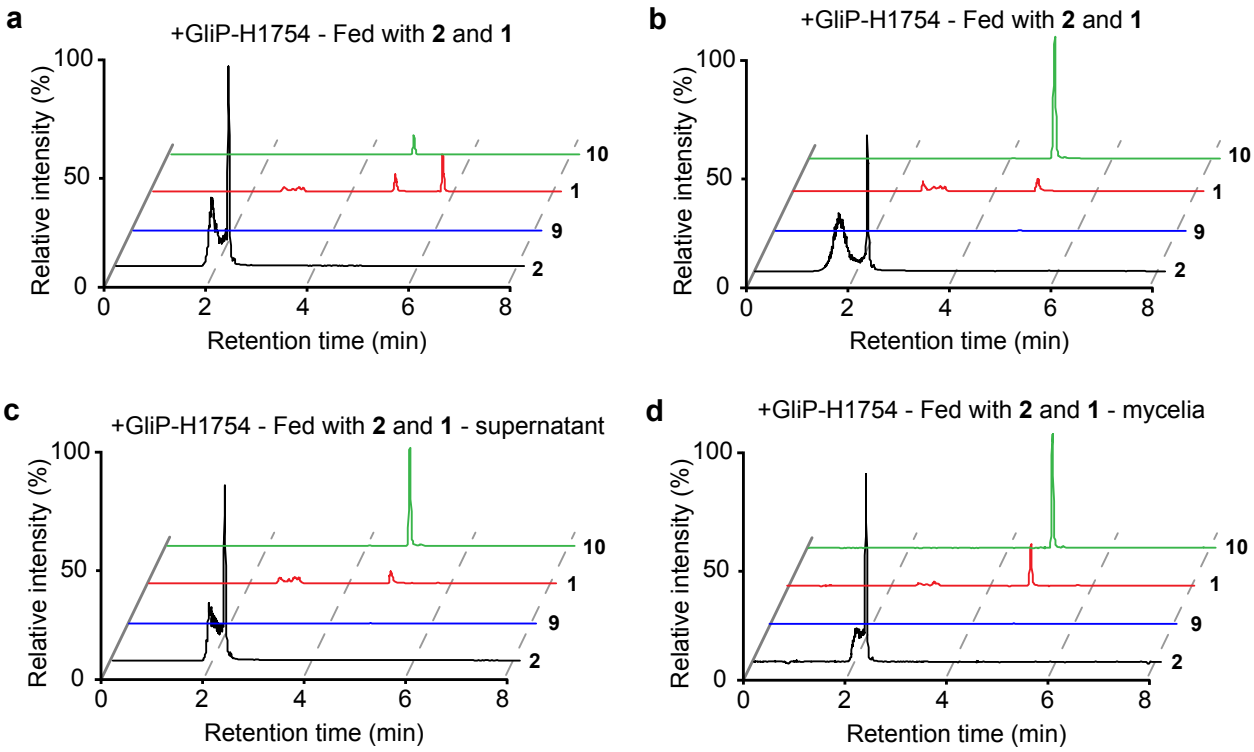


Figure 2.9. LC-HRMS analysis of whole metabolome extracts from GliP-H1754A mutant *A. fumigatus* strains supplemented with exogenous 1 and 2 during growth. (**a-d**) Overlaid extracted-ion-chromatograms for **1** (red lines), **2** (black lines), **9** (blue lines) and **10** (green lines).

.

One possible explanation for the absence of more highly modified *gli* pathway metabolites in this feeding experiment is that **2** did not get in to the cell. In order to determine if **2** was incorporated by the fungus, we filtered the mycelia from the culture broth of a 25 mL liquid culture (see **Appendix B** for details), then washed the mycelia twice. The washes and supernatant were extracted separately from the mycelia, and whole-metabolome LC-HRMS analysis of these samples revealed that **2** was present in the mycelia as well as the supernatant (**Figure 2.9c-d**). Therefore, if **2** was an intermediate in the *gli* pathway and a bona fide substrate for *gli* cluster tailoring enzymes, we should have observed additional *gli* pathway metabolites, however we did not.

Collectively these results strongly suggest that **2** is not the substrate for at least some of *gli* cluster tailoring enzymes.

GliJ dipeptidase cleaves L-Phe-L-Ser: The *gli* cluster dipeptidase, GliJ, was previously shown to be responsible for converting cyclo(L-Phe-L-Ser)-Cys-Gly (**5**) into cyclo(L-Phe-L-Ser)-Cys (**7**), which is then converted into dithiol-cyclo(L-Phe-L-Ser) leaving only one remaining step to form dithio-cyclo(L-Phe-L-Ser) by GliT[17,26]. The initial publication on GliJ did not provide information on substrate specificity and the enzymatic activity was not characterized by itself, but rather shown through a complex *in vitro* assay containing two additional enzymes, GliK and GliI[17]. This doesn't leave any doubt as to the role of GliJ in the anabolism of **1**, however as we have previously seen, enzymes in BGCs occasionally perform multiple functions[18]. In order to determine if GliJ could have a dual-function in a manner similar to HasE (described in Chapter 1), we purchased the gene for GliJ, then expressed and purified H_6-tagged GliJ from *E. coli* (see **Appendix B**). We tested GliJ activity on a range of dipeptide substrates and monitored the hydrolytic activity by LC-MS (**Figure B12**). In addition to cleaving thiol-capped L-Cys-L-Gly, GliJ also readily cleaved L-Phe-L-Ser and several other dipeptides (**Figure B12**). Unfortunately, we could not compare GliJ activity on L-Phe-L-Ser to **5**, as this intermediate would be very difficult to access synthetically. The first publication on GliJ used 260 L of a ΔGliK *A. fumigatus* strain to isolate **4**, which was then processed by recombinant GliK prior to showing GliJ activity (**Figure 2.2**)[17]. This would be a time-intensive experiment to repeat in order to show the difference in turnover rates between L-Phe-L-Ser and what has been hypothesized to be the true substrate for GliJ, **4**.

Conclusions: Our reentry into the *gli* pathway has revealed that despite extensive speculation[12,15-17,24], the cyclization of L-Phe-L-Ser en route to gliotoxin is not a spontaneous event but rather requires the cyclization activity of the C_2 domain. We also provide evidence for the function of the enigmatic terminal thiolation domain of GliP, where it serves as a location for tethered L-Phe-L-Ser to be modified by *gli* cluster tailoring enzymes. The GliP C_2 domain and similar domains found in other homologous NRPSs, should be re-annotated as C_T domains from this point forward. In the future it will be necessary to characterize the early modifications of L-Phe-L-Ser by other *gli* cluster tailoring enzymes. In particular, *in vitro* reconstitution of the activity for the oxidases GliC and GliF would provide key evidence to clarify the full mechanism of gliotoxin biosynthesis. This will be important because gliotoxin is representative of a large class of NRPS-derived 2,5-diketopiperazines that are likely produced via similar biosynthetic mechanisms.

REFERENCES

1. Johnson, J. R., Bruce, W. F. & Dutcher, J. D. Gliotoxin, The Antibiotic Principle of Gliocladium fimbriatum. I. Production, Physical and Biological Properties. *J Am Chem Soc* **65,** 2005–2009, (1943).

2. Scharf, D. H., Brakhage, A. A. & Mukherjee, P. K. Gliotoxin--bane or boon? *Environ Microbiol* **18,** 1096–1109, (2016).

3. Fridrichsons, J. & McL Mathieson, A. The crystal structure of gliotoxin. *Acta Cryst* **23,** 439–448, (1967).

4. Johnson, J. R., Wildi, B. S. & Woodward, R. B. The structure of gliotoxin. *J Am Chem Soc* **4**, 1001, (1958).

5. Fukuyama, T. & Kishi, Y. A total synthesis of gliotoxin. *J Am Chem Soc* **21**, 6723-6724, (1976).

6. Fukuyama, T., Nakatsuka, S.-I. & Kishi, Y. Total synthesis of gliotoxin, dehydrogliotoxin and hyalodendrin. *Tetrahedron* **37,** 2045–2078, (1981).

7. Waring, P. & Beaver, J. Gliotoxin and related epipolythiodioxopiperazines. *General Pharmacology: Vascul Sys* **27,** 1311–1316, (1996).

8. Waring, P. DNA fragmentation induced in macrophages by gliotoxin does not require protein synthesis and is preceded by raised inositol triphosphate levels. *J Biol Chem* **24**, 14476-14480, (1990).

9. Denning, D. W. Invasive aspergillosis. *Clinic Infect Dis* **4**, 781-803, (1998).

10. Sutton, P. & Waring, P. Exacerbation of invasive aspergillosis by the immunosuppressive fungal metabolite, gliotoxin. *Immun Cell Biol* **74**, 318-322, (1996).

11. Fosker, N., Fraser, A., García, J. L., García, M. J. & Goble, A. Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus. *Nature* **438**, 1151-1156, (2005).

12. Dolan, S. K., O'Keeffe, G., Jones, G. W. & Doyle, S. Resistance is not futile: gliotoxin biosynthesis, functionality and utility. *Trends Microbiol* **23,** 419–42, (2015).

13. Balibar, C. J. & Walsh, C. T. GliP, a Multimodular Nonribosomal Peptide Synthetase in Aspergillus fumigatus, Makes the Diketopiperazine Scaffold of Gliotoxin. *Biochemistry* **45,** 15029–15038, (2006).

14. Forseth, R., Fox, E. & Chung, D. Identification of Cryptic Products of the Gliotoxin Gene Cluster using NMR-based Comparative Metabolomics and a Model for Gligotoxin Biosynthesis. *J Am Chem Soc* **133**, 9678-9681, (2011).

15. Scharf, D. H. *et al.* Biosynthesis and function of gliotoxin in Aspergillus fumigatus. *Appl Microbiol Biotechnol* **93,** 467–472, (2011).

16. Scharf, D. H. *et al.* A Dedicated Glutathione S-Transferase Mediates Carbon–Sulfur Bond Formation in Gliotoxin Biosynthesis. *J Am Chem Soc* **133,** 12322–12325, (2011).

17. Scharf, D. H. *et al.* Epidithiodiketopiperazine Biosynthesis: A Four-Enzyme Cascade Converts Glutathione Conjugates into Transannular Disulfide Bridges. *Angew Chem Int Ed* **52,** 11092–11095, (2013).

18. Yin, W.-B. *et al.* A Nonribosomal Peptide Synthetase-Derived Iron(III) Complex from the Pathogenic Fungus Aspergillus fumigatus. *J Am Chem Soc* **135,** 2064–2067, (2013).

19. Gao, X. *et al.* Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. *Nat Chem Biol* **8**, 823-830, (2012).

20. Gao, X., Chooi, Y. H., Ames, B. D. & Wang, P. Fungal indole alkaloid biosynthesis: genetic and biochemical investigation of the tryptoquialanine pathway in Penicillium aethiopicum. *J Am Chem Soc* **133**, 2729-2741, (2011).

21. Kremer, A., Westrich, L. & Li, S.-M. A 7-dimethylallyltryptophan synthase from Aspergillus fumigatus: overproduction, purification and biochemical characterization. *Microbiology* **153,** 3409–3416, (2007).

22. Cramer, R. A. *et al.* Disruption of a nonribosomal peptide synthetase in Aspergillus fumigatus eliminates gliotoxin production. *Eukaryotic Cell* **5,** 972–980, (2006).

23. Forseth, R. R. *et al.* Homologous NRPS-like Gene Clusters Mediate Redundant Small-Molecule Biosynthesis in Aspergillus flavus. *Angew Chem Int Ed* **52,** 1590–1594, (2012).

24. Chang, S.-L., Chiang, Y.-M., Yeh, H.-H., Wu, T.-K. & Wang, C. C. C. Reconstitution of the early steps of gliotoxin biosynthesis in Aspergillus nidulans reveals the role of the monooxygenase GliC. *Bioorg Med Chem Lett* **23,** 2155–2157, (2013).

25. Forseth, R. R. *et al.* Identification of cryptic products of the gliotoxin gene cluster

using NMR-based comparative metabolomics and a model for gliotoxin biosynthesis. *J Am Chem Soc* **133,** 9678–9681, (2011).

26. Schrettl, M. *et al.* Self-protection against gliotoxin--a component of the gliotoxin biosynthetic cluster, GliT, completely protects Aspergillus fumigatus against exogenous gliotoxin. *PLoS Pathog* **6,** e1000952, (2010).

CHAPTER 3

PLANT-LIKE ISOQUINOLINE ALKALOID BIOSYNTHESIS IN *ASPERGILLUS FUMIGATUS*

Abstract: Natural product discovery efforts have focused primarily on microbial biosynthetic gene clusters (BGCs) containing large multi-modular PKSs and NRPSs; however, sequencing of fungal genomes has revealed a vast number of BGCs containing smaller NRPS-*like* genes of unknown biosynthetic function. Using comparative metabolomics, we show that a BGC in the human pathogen *Aspergillus fumigatus* named *fsq*, which contains an NRPS-like gene lacking a condensation domain, produces several novel isoquinoline alkaloids, the fumisoquins. These compounds derive from carbon-carbon bond formation between two amino acid-derived moieties followed by a sequence that is directly analogous to isoquinoline alkaloid biosynthesis in plants. Fumisoquin biosynthesis requires the *N*-methyltransferase FsqC and the FAD-dependent oxidase FsqB, which represent functional analogs of coclaurine *N*-methyl transferase and berberine bridge enzyme in plants. Our results show that BGCs containing incomplete NRPS modules may reveal new biosynthetic paradigms and suggest that plant-like isoquinoline biosynthesis occurs in diverse fungi.

Introduction: One major source of medicinally important small molecules are microbial BGCs that encode polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs), two classes of multi-functional mega-enzymes that are usually accompanied by sets of tailoring enzymes. The recent surge in the sequencing of fungal genomes has revealed large numbers of BGCs that do not appear to encode enzymes involved in the production of any known metabolites ("orphan BGCs") and thus may harbor new biosynthetic capabilities[1-3].

In search of new biosynthetic mechanisms, we focused on orphan BGCs that contain small NRPS-like genes that diverge from canonical NRPSs in their domain structure. Canonical NRPSs include, at a minimum, one adenylation domain that selects and activates an amino acid (or related building block), one thiolation domain for covalent attachment of the activated amino acid, and a condensation domain that catalyzes formation of a peptide bond between the tethered amino acid and another substrate[4-6]. However, sequencing of fungal genomes has additionally revealed numerous genes that encode only a subset of these domains[1,2]. Such non-canonical, NRPS-like genes may feature adenylation and thiolation domains, but lack condensation domains, and thus likely do not function as peptide synthetases[7,8]. We hypothesized that analysis of BGCs including small NRPS-like genes encoding incomplete modules would likely reveal novel biosynthetic functions.

Here we show that the fsq gene cluster in A. fumigatus, which features an NRPS-like gene, fsqF, produces a series of novel isoquinoline alkaloids, the fumisoquins. We then demonstrate that FsqF, which lacks a canonical condensation domain, is required for carbon-carbon bond formation between L-serine and L-tyrosine-derived building

blocks in the fumisoquin biosynthetic pathway. Two additional enzymes encoded by the fsq cluster, the N-methyltransferase FsqC and the FAD-dependent oxidase FsqB, catalyze the formation of the isoquinoline ring system in the fumisoquins via a sequence that we show is directly analogous to the biosynthesis of a prominent group of plant isoquinoline alkaloids, one of the largest families of medicinally important natural products[9,10].

***FsqF* encodes an incomplete NRPS module:** Expression of many fungal orphan BGCs is under the control of the nuclear protein LaeA, a global regulator of morphogenesis and virulence factor in *A. fumigatus* and other pathogenic fungi[11,12]. Among LaeA-regulated orphan BGCs in the human pathogen *A. fumigatus*, the *fsq* cluster features a small NRPS-like gene, *fsqF*, which encodes an adenylation, thiolation, reductase and pyridoxal-phosphate dependent aminotransferase domain (**Figure 3.1a**). Notably, FsqF lacks the condensation domain that is essential for canonical peptide bond formation, and bioinformatic analysis of the vicinity of *fsqF* did not reveal any genes that could code for a second NRPS or free-standing condensation domain[13]. We found that expression of *fsqF* is regulated by the $Zn(II)_2Cys_6$-type transcriptional activator FsqA (**Figure 3.1a** and **see Apendix C, Figure C1**). $Zn(II)_2Cys_6$ transcription factors are unique to fungi and have been shown to regulate many metabolic pathways, e.g. aflatoxin biosynthesis in *A. flavus*[14]. In addition to *fsqF*, the transcription factor FsqA was found to regulate expression of five adjacent genes, annotated as ABC transporter (*fsqE*), fructosyl amino acid oxidase (*fsqB*), *N*-methyl-transferase (*fsqC*), phenol 2-monooxygenase (*fsqG*), and ATP-grasp enzyme (*fsqD*), defining the boundaries of the *fsq* cluster (**Figure C1**)[13].

Identification of *fsq*-cluster derived metabolites: Overexpression of *fsqA* was associated with accumulation of a characteristic brown pigment in the surrounding media (**Figure C2**), suggesting secretion of a colored metabolite(s). For the identification of *fsq*-derived metabolites we employed 2D NMR-based comparative metabolomics (differential analysis by 2D-NMR spectroscopy, DANS), which facilitates comprehensive and largely unbiased detection of new metabolites associated with gene cluster overexpression[15,16]. DANS-based comparison of whole metabolome extracts from overexpression (OE::*fsqA*) and deletion (Δ*fsqA*) strains revealed several differential signals, most prominently two sets of OE-dependent doublets between 4.6 and 5.0 ppm with large coupling constants of 18-20 Hz (**Figure 3.1b**), suggesting a diastereotopic benzylic CH_2 group within a bi- or oligocyclic structure. The OE::*fsqA*-dependent metabolites were characterized further following partial purification via reversed-phase chromatography, using a standard suite of 2D NMR spectra and high-resolution mass spectrometry (HRMS) (**Table C6-12**). These analyses revealed two novel tricyclic isoquinoline derivatives, the pyrido[1,2-

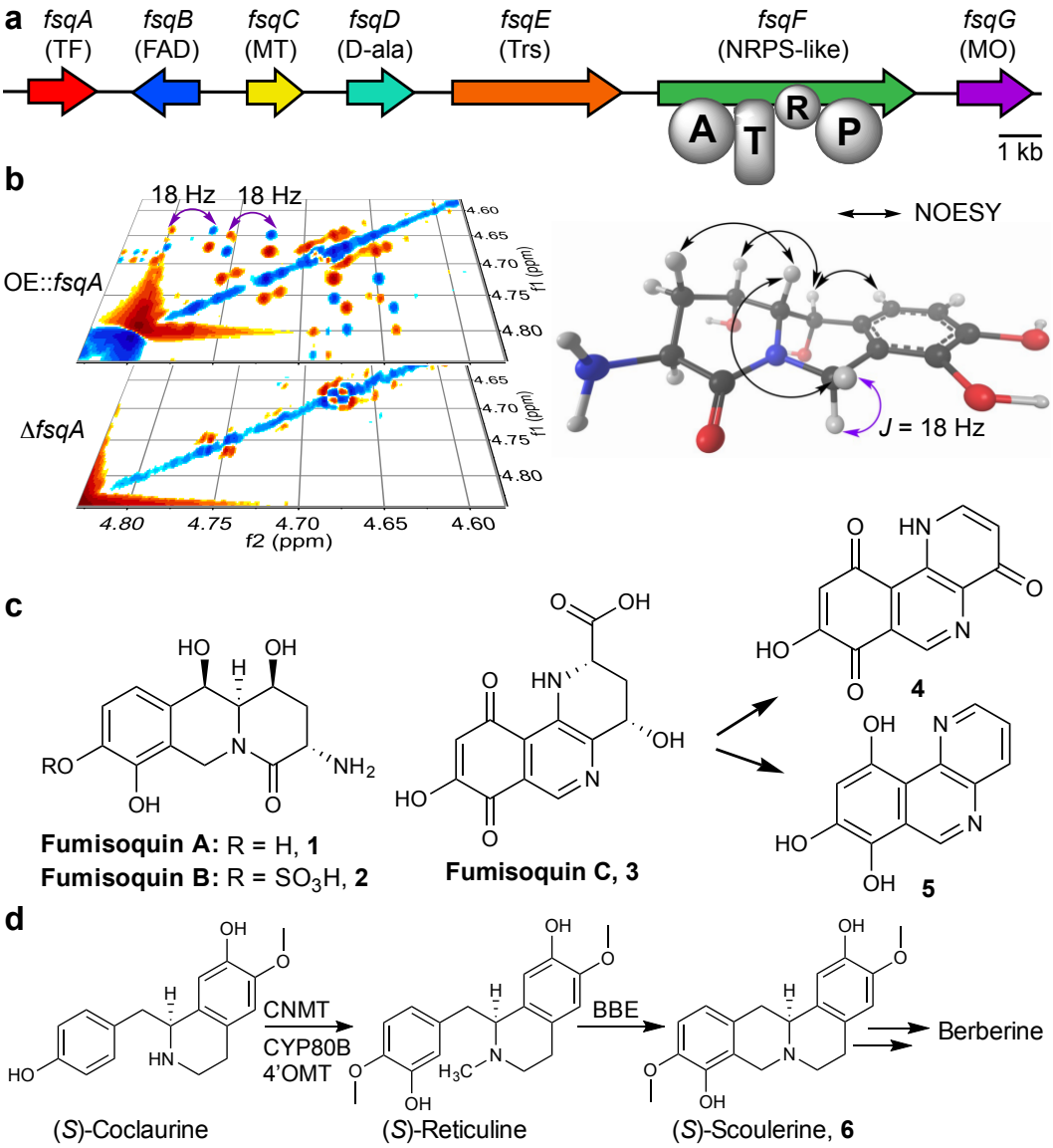


Figure 3.1. Analysis of the *fsq* gene cluster and metabolite production. (**a**) *fsq* gene cluster and putative assignments of encoded proteins. TF: transcription factor; FAD: FAD-binding domain protein; MT: *N*-methyltransferase; D-ala: ATP-grasp enzyme (D-alanine ligase); Trs: transporter; NRPS: nonribosomal peptide synthetase (A = adenylation, T = thiolation, R = short-chain dehydrogenase/reductase domain, P: pyridoxal phosphate binding domain); MO: phenol 2-monooxygenase. (**b**) Section of the dqfCOSY spectra of OE::*fsqA* and Δ*fsqA* metabolite extracts used for comparative metabolomics (DANS). (**c**) Identified *fsq*-dependent compounds fumisoquin A and B (**1** and **2**), including stereochemical assignments via NOESY for **1** (see **Tables C6 and C7**), as well as structure of fumisoquin C (**3**), which decomposes to **4** and **5**. (**d**) Biosynthesis of structurally related isoquinoline alkaloids in plants via coclaurine *N*-methyltransferase (CNMT), *N*-methylcoclaurine 3'-monooxygenase (CYP80B), 3'-hydroxy-*N*-methyl-(*S*)-coclaurine 4'-*O*-methyltransferase (4'OMT) and berberine bridge enzyme (BBE).

b]isoquinolines, fumisoquin A (**1**, **Figure 3.1c**) and fumisoquin B (**2**, **Figure 3.1c**) as major metabolites associated with *fsq*-overexpression. Fumisoquin A and B decompose gradually during chromatography, likely as a result of the oxidation-prone hydroquinone moiety. We further noted a deeply purple OE-dependent metabolite that was not captured by DANS because of decomposition during sample preparation. Optimization of extraction conditions and reverse-phase fractionation, followed by 2D NMR spectroscopic and HRMS analysis allowed to identify a third isoquinoline, fumisoquin C (**3**, **Figure 3.1c**), as the deeply purple colored metabolite (**Figure C3 and Table C3**). Upon standing or during chromatography, fumisoquin C decomposes into compounds 4 and 5, which are more stable and were isolated (**Figure 3.1c**, and **Figure C4a**). Treatment of partially purified fumisoquin C with (trimethylsilyl)diazomethane furnished the corresponding dimethyl derivative (**Figure C4b**), which is much less prone to decomposition and provided additional NMR spectroscopic data to confirm structure and relative configuration of fumisoquin C (Supplementary Note).

A plant-analogous isoquinoline formation pathway: Among known fungal natural products, isoquinoline alkaloids are rare, and the tricyclic ring systems of the fumisoquins, to the best of our knowledge, are unprecedented from fungi. However, the identified dihydroxylated isoquinolines are strikingly reminiscent of a large and diverse family of plant-derived isoquinoline alkaloids (pyrido[1,2-*b*]isoquinolines, e.g. Scoulerine, **6**, **Figure 3.1d**), whose biosynthetic pathways have been studied extensively[17-19]. Plant pyrido[1,2-*b*]isoquinoline biosynthesis proceeds via *N*-methylation of a tyrosine-derived catechol precursor, followed by oxidative cyclization via the FAD-dependent berberine bridge

enzyme (BBE, **Figure 3.1d**)[17], a member of a large group of flavin-dependent amine oxidases[20,21]. The cyclization is believed to proceed via a two-step process, beginning with FAD-dependent oxidation of the methyl group to an iminium species followed by electrophilic attack on the deprotonated phenol[17,19]. Because of the manifold

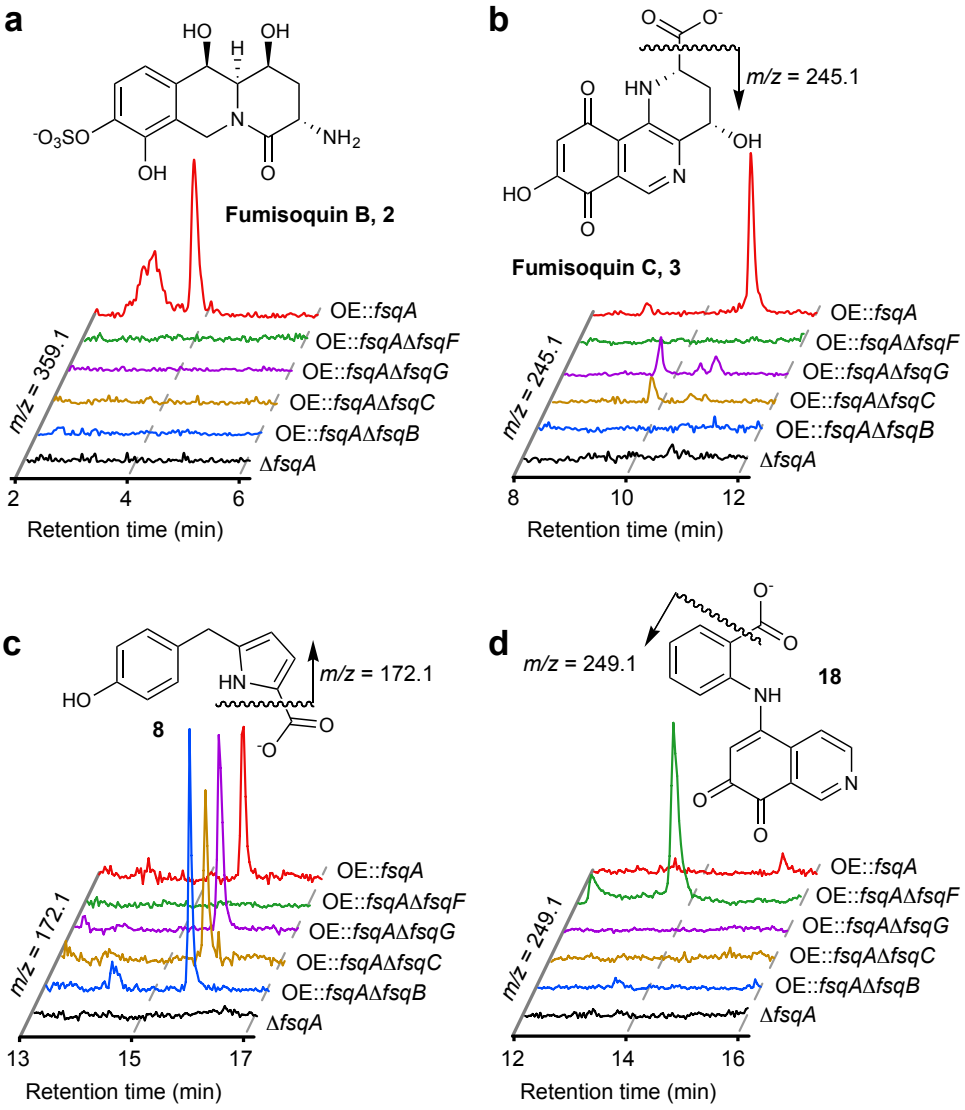


Figure 3.2. Comparative metabolomics of *fsq* gene cluster mutants by HPLC-MS. (**a-d**) HPLC-MS analysis of *fsq*-gene deletion strains in OE::*fsqA* background, shows ion chromatograms representing major products and shunt metabolites.

pharmaceutical uses of plant isoquinoline alkaloids, there is considerable interest in developing microbial-based production approaches[22-24], and yeast-based expression systems for several groups of plant isoquinoline alkaloids have been developed, including the berberines[25-30]. We noted that the *fsq* cluster features a set of genes that may encode functional analogs of the berberine biosynthetic enzymes, including the putative phenol 2-monooxygenase FsqG, *N*-methyltransferase FsqC, and FAD-dependent oxidase FsqB. Bioinformatic analysis revealed several orphan gene clusters in fungi that feature homologous sets of genes (**Figure C5**). Moreover, the only other known fungal isoquinolines have been linked to a gene cluster containing homologs for these three genes[31].

To investigate the biosynthesis of the fumisoquins, we created deletion strains in the OE::*fsqA* background for all genes in the *fsq* cluster, except for *fsqD* and the putative transporter *fsqE*, which we were unable to delete, which could be due to a protective function of these proteins as has been found in several other fungal BGCs[1]. We then compared the metabolomes of each of these deletion strains with those of OE::*fsqA* and Δ*fsqA* via DANS and HPLC-MS (**Figure 3.2a-d**). We found that deletion of the *N*-methyltransferase FsqC, phenol 2-monooxygenase FsqG, or FAD-dependent oxidase FsqB in OE::*fsqA* background led to complete abolishment of isoquinoline alkaloid production and accumulation of a series of benzyl pyrroles (**7-8**, **Figure 3.2c and Figure 3.3**), including the previously reported fumipyrrole, 7 (where the *fsq* cluster was referred to as *fmp*[32]), smaller amounts of which were also produced in the OE::*fsqA* control strain (**Figure 3.2c**). Comparison of the carbon skeletons of the pyrroles (**7-8**, **Figure 3.3**) with those of the fumisoquins provided support for the predicted *N*-methyltransferase and

phenol 2-monooxygenase functions of FsqC and FsqG, respectively, and suggested that the fumisoquins form from an open-chain precursor such as 9 (**Figure 3.3**), in a manner very similar to that of plant benzophenanthridine alkaloid biosynthesis[18].

The structure of the putative open-chain precursor **9** suggested that it may be derived from tyrosine and a dehydroalanine equivalent (**10**, **Figure 3.3**), which could be produced from serine or cysteine via the pyridoxal-phosphate dependent aminotransferase domain of the NRPS-like protein FsqF[10,33]. Growing the OE::*fsqA* strain in media supplemented with different stable-isotope labeled amino acids, we found that L-tyrosine and L-serine are incorporated into the fumisoquins and pyrroles (**1-3** and **8**) in a manner consistent with the biosynthetic model shown in **Figure 3.3**, whereas L-cysteine is not incorporated (**Figure 3.4a,b** and **Figure C6**, **C7b-d**). Furthermore, we showed that the L-methionine methyl group is incorporated into the fumisoquin ring systems at the predicted position, in agreement with the putative function of FsqC as an *S*-adenosyl methionine-dependent *N*-methyltransferase (**Figure 3.4c** and **Figure C8**).

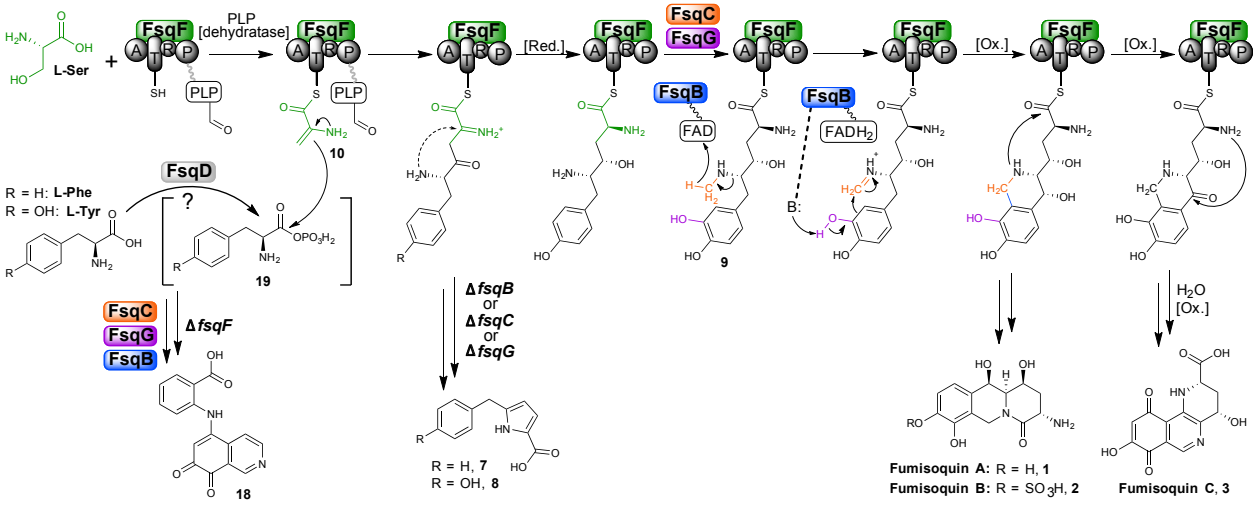


Figure 3.3. Model for fumisoquin biosynthesis. (A = adenylation, T = thiolation, R = short-chain dehydrogenase/reductase domain, P: pyridoxal phosphate binding domain), [Red.] = reduction, [Ox.] = oxidation. The proposed mechanism for FsqB-mediated isoquinoline formation is based on functional similarity to plant BBE[17-19], see **Figure 3.5** for details. The role of the ATP-grasp enzyme FsqD in activating tyrosine to form **19** is putative.

FsqB is a functional analog of plant BBE: To test whether FsqB catalyzes isoquinoline ring formation in **1-3** and thus acts as a functional ortholog of plant BBE, we recombinantly produced polyhistidine-tagged FsqB in *E. coli* (**Figure C9**). We then assayed its activity on a range of potential substrates that mimicked the tyrosine-derived portion of putative intermediate 9. Whereas *N*-methyl dopamine did not react, we found that FsqB converts *N*-methyl DOPA 11, prepared via a simple one-step procedure, directly into isoquinoline 12 (**Figure 3.5a** and **Figure C10**, **C12a**). We first tested FsqB activity on **11** on a small scale suitable for LCMS, which yielded near quantitative conversion to **12** within 3–4 h (100 mM phosphate buffer, pH 7, K_M = 142.5 ± 42.9 μM) (**Figure 3.5b** and **Figure C11**). Next we demonstrated isoquinoline formation from **11** on a milligram scale by conducting the reaction in an NMR tube using deuterated phosphate buffer (**Figure 3.5c** and see

Appendix C). Notably, FsqB-catalyzed formation of **12** proceeds with complete regioselectivity, as the alternative cyclization product, 6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, **13**, which does form alongside **12** in a non-enzymatic reaction of DOPA with formaldehyde, was not observed (**Figure C10** and **Appendix C**). In contrast, we observed no cyclized products when incubating FsqB with *N*-methyl-L-tyrosine, *N,N*-dimethyl DOPA, or the BBE-substrate (*S*)-reticuline, suggesting that both a catechol moiety and a secondary β-*N*-methylamine are required for FsqB activity (**Figure C12b-d**).

Next we asked whether the mechanism of isoquinoline formation as catalyzed by FsqB is directly analogous to that of the extensively studied plant enzyme, BBE[9,17,19,34]. Fluorescence analysis of recombinant FsqB showed excitation and emission maxima at 461 nm and 530 nm, respectively (**Figure C9**), suggesting FsqB utilized a flavin cofactor, as predicted[35]. Further proteomic analysis of FsqB confirmed the presence of FAD, which was found to be covalently bound to cysteine-414 (**Figure C13a** and **Table C3**). FAD is also part of BBE where it is bound to a histidine residue[17]. Flavin reduction during catalytic cycling of BBE is recovered with molecular oxygen generating hydrogen peroxide[17], a characteristic that was also observed with FsqB utilizing an Amplex Red assay (**Figure C13b**). Isoquinoline formation as catalyzed by BBE likely proceeds via an imine intermediate in a two-step mechanism[19]. To test whether isoquinoline formation via FsqB also involves an imine intermediate, we incubated FsqB with *N*-methyl-L-tyrosine **14**, which we hypothesized may represent an oxidation-competent substrate for FsqB but, as shown above, does not get cyclized, in the presence of the electrophile trapping agent dimedone, **15**. This reaction produced the adduct **16** as well as free tyrosine, consistent

with capture of an imine intermediate, **17**, by dimedone or water (**Figure 3.5d**). Therefore, it appears that the fungal enzyme FsqB catalyzes isoquinoline formation via a two-step mechanism that is analogous to that of plant-derived BBE.

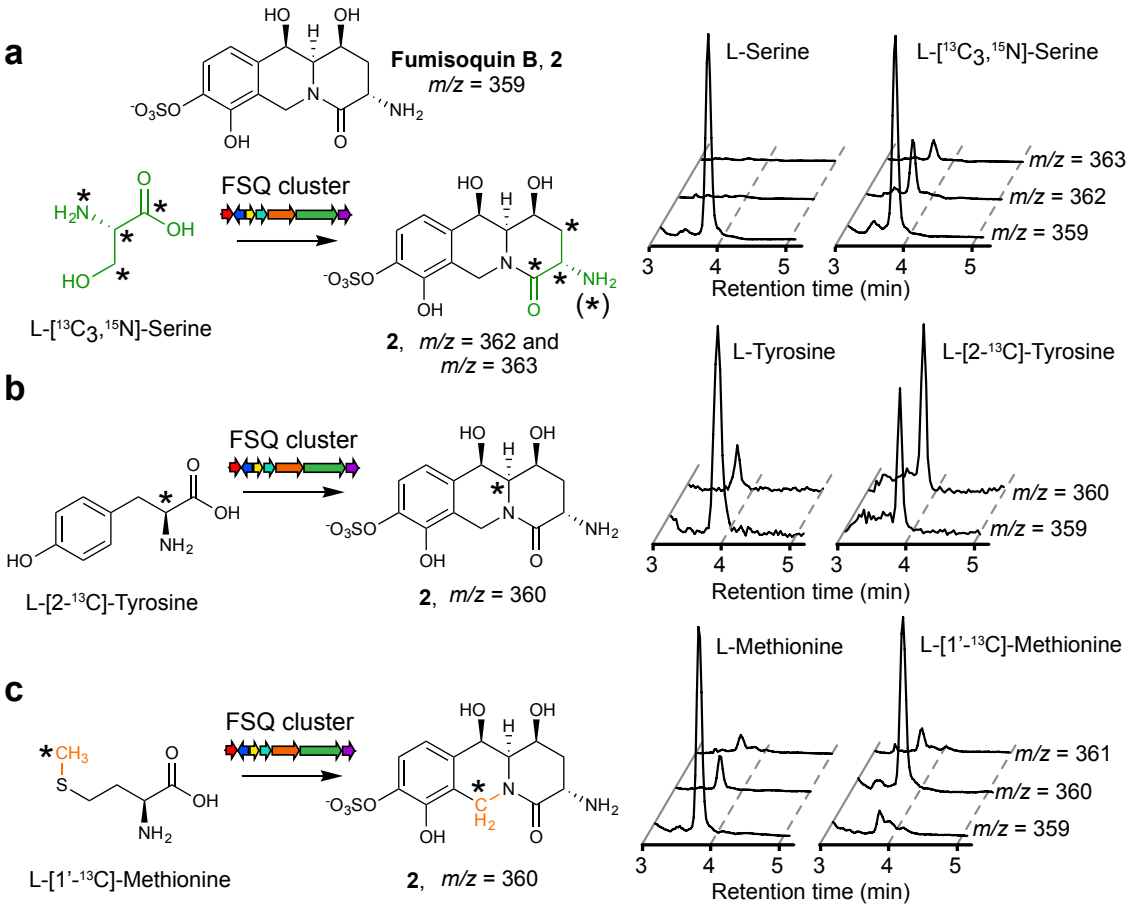


Figure 3.4. The fumisoquins incorporate L-serine, L-tyrosine, and an L-methionine-derived methyl group. Shown are HPLC-MS ion chromatograms derived from fungal cultures grown with (**a**) L-serine or L-[$^{13}C_3$,^{15}N]-serine, (**b**) L-tyrosine or L-[2-^{13}C]-tyrosine and (**c**) L-methionine or L-[1'-^{13}C]-methionine. In contrast, L-cysteine is not incorporated (**Figure C6**). See **Figure C7** for mass spectra. (*) denotes labeled isotopes. In the case of L-[$^{13}C_3$,^{15}N]-serine, loss of ^{15}N due to partial deamination/amination contributes to increased m/z = 362.

Carbon-carbon bond formation requires FsqF: Next we recombinantly produced the FsqF adenylation domain to test whether it activates L-tyrosine or L-serine, as suggested by the stable isotope labeling study, using the ATP-[^{32}P]pyrophosphate exchange activity[36]. However, neither enantiomer of serine or tyrosine showed any activity in the *in vitro* assay nor did any other tested standard amino acid (**Figure C14**). These results may indicate that the recombinantly expressed FsqF adenylation domain was not functional or that the true substrate is not a standard amino acid. Based on the homology of the PLP-dependent domain in FsqF to other serine dehydrogenases[13,33], we hypothesized that L-serine may first be converted into dehydroalanine, which is then activated and tethered to the thiolation domain (**Figure 3.3**).

To further investigate the role of the NRPS-like enzyme FsqF in fumisoquin biosynthesis, we analyzed the effect of *fsqF* deletion on the OE::*fsqA* metabolome. As anticipated, production of the fumisoquins (**1-3**) as well as the pyrroles (**7-8**) was completely abolished in OE::*fsqA*-Δ*fsqF* (**Figure 3.2d**). Significantly, we also observed production of a single, brightly red shunt metabolite, the anthranilic acid-substituted isoquinoline **18** (**Figure 3.2d** and **Figure 3.3**). Feeding experiments of the OE::*fsqA*-Δ*fsqF* mutant with stable-isotope labeled L-tyrosine confirmed that the isoquinoline ring system in **18** is derived from tyrosine, as in the case of the fumisoquins (**Figure C15**). Incubation of **12**, obtained *in vitro* from recombinant FsqB, with anthranilic acid (**19**) led to formation of shunt metabolite **18**, consistent with a model in which shunt metabolite **18** forms non-enzymatically from **12** via oxidation to the corresponding *o*-quinone, subsequent capture of anthranilic acid, and decarboxylation (**Figure C16**). Compound **12** was not observed in any of the analyzed OE::*fsqA*-Δ*fsqF* extracts, likely as a result of this

compound's high susceptibility to oxidize to the reactive *o*-quinone, which is rapidly captured by anthranilic acid (**Figure C16**).

The identification of shunt metabolite **18** indicates that the NRPS-like protein FsqF is only required for addition of a serine-derived dehydroalanine moiety to activated tyrosine, but is not essential for the subsequent steps leading to isoquinoline formation, and that a different enzyme, likely the ATP-grasp enzyme FsqD, is responsible for activation of tyrosine (**Figure 3.3**). The production of shunt metabolite 7 would then indicate that FsqD is also capable of activating phenylalanine; however, phenylalanine-

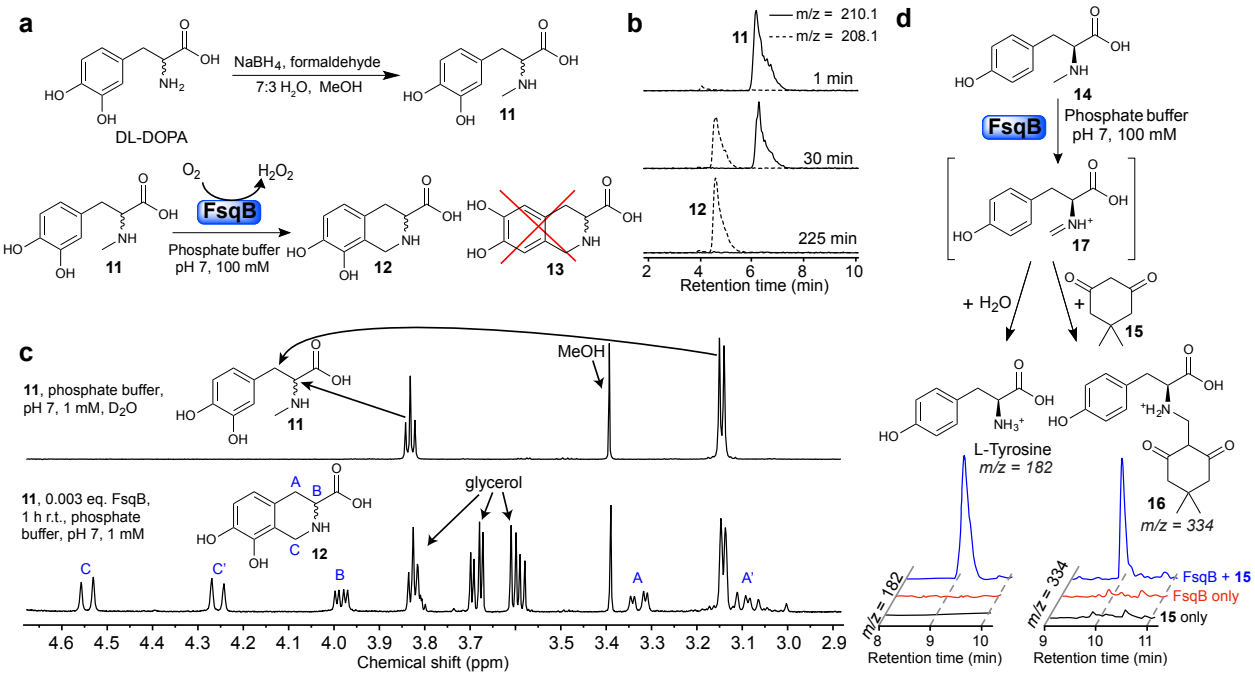


Figure 3.5. Enzymatic activity of recombinant FsqB. (**a**) Synthesis of *N*-methyl-3',4'-dihydroxy-DL-phenylalanine **11** (top) and regioselective conversion of **11** into **12** by purified recombinant FsqB (bottom). The isomer **13** does not form. (**b**) Ion chromatograms at indicated timepoints for the FsqB-catalyzed *in vitro* conversion of **11** to **12** (100 mM phoshate buffer, pH 7, 100:1 substrate to enzyme). (**c**) ^{1}H NMR spectra of a sample of **11** (1 mM phosphate buffer in D_2O) prior to FsqB addition (top) and 1 h post FsqB addition (333:1 substrate to enzyme), showing selective conversion of **11** into **12** (bottom) (**d**) Ion chromatograms showing formation of tyrosine and **16** as a result of capture of intermediate **17** by H_2O or dimedone, **15** (1.5 µM FsqB and 400 µM *N*-methyl-L-tyrosine in 100 mM phoshate buffer, pH 7).

derived intermediates only lead to the production of **7**, as expected based on the predicted function of FsqG as a phenol 2-monooxygenase[13] and given the observation that recombinantly produced FsqB requires a catechol as substrate. FsqD has homology to carboxylate-amine ligases that furnish aminoacyl phosphate from ATP and amino acid precursors[37,38]. In a recent example, the FsqD homolog PGM1 was shown to activate non-proteinogenic amino acids for peptide-bond formation in pheganomycin biosynthesis[39]. In contrast, FsqD appears to activate tyrosine (or phenylalanine) for subsequent condensation with serine-derived dehydroalanine (**Figure 3.3**), providing a first example for a novel carbon-carbon bond forming strategy in fungi.

Conclusions: Our analysis of the *fsq* cluster revealed a plant-like strategy for the biosynthesis of oligocyclic alkaloids in fungi. The biosynthetic steps of phenol-hydroxylation, *N*-methylation, and oxidative cyclization appears to be the same in plants and fungi, and thus the corresponding enzymes can be considered functionally equivalent. Sequence analysis does not suggest homology for these proteins, though it is notable that cyclization proceeds via analogous two-step mechanisms that are catalyzed by FAD-dependent oxidases in both fungi and plants[19], presenting a striking case of convergent evolution of specialized metabolic pathways. Given that analysis of available fungal genomes[13] revealed many co-occurrences of homologs of the *N*-methyltransferase *fsqC* and FAD-dependent oxidase *fsqB* (**Figure C5**), it seems likely that fungi are capable of producing a diverse range of yet undiscovered isoquinoline

alkaloids. For example, in addition to *Aspergillus* spp., the genomes of plant pathogenic *Fusarium* spp. feature BGCs that include homologues for most *fsq* genes[40].

Together with other recent examples[7,8], the identification of the fumisoquins shows that analysis of gene clusters containing incomplete NRPS modules can reveal intriguing new structures and biosynthetic strategies. It remains unclear, however, whether the biosynthetic roles of NRPS-like enzymes can be predicted, as current examples, though few, hint at considerable functional diversity. For example, two NRPS-like enzymes in *A. flavus*, named LnaA and LnaB, have been shown to produce a series of tyrosine-derived piperazines, pyridines and morpholines[7], whereas an NRPS-like enzyme from *A. terreus* was shown to reduce products of an accompanying non-reducing PKS[8]. The example of the *fsq* pathway demonstrates that NRPS-like enzymes lacking condensation domains may nonetheless contribute to ligating amino acids, albeit via formation of carbon-carbon bonds. Characterization of BGCs featuring small NRPS-like genes may thus reveal new types of alliances of NRPS-like enzymes with other amino acid-activating proteins, e.g. ATP-grasp enzymes such as FsqD.

This work supports a mining strategy of assessing fungal natural products, beyond not only canonical NRPS but also PKS-like genes and even those fungal clusters lacking NRPS or PKS-like domain genes all together. Recent examples have focused on cyclic ribosomal peptides[41] but a bioinformatic scan of fungal genomes and expression data shows the existence of yet more alternative biosynthetic clusters – lacking NRPS-like, PKS-like and canonical ribosomal peptides - suggestive of unexplored chemical diversity[42]. Efforts to characterize such alternative clusters may yield even more unique fungal chemistry.

REFERENCES

1. Keller, N. P. Translating biosynthetic gene clusters into fungal armor and weaponry. *Nat Chem Biol* **11**, 671-677, (2015).

2. Sanchez, J. F., Somoza, A. D., Keller, N. P. & Wang, C. C. Advances in Aspergillus secondary metabolite research in the post-genomic era. *Nat Prod Rep* **29**, 351-371, (2012).

3. Brakhage, A. A. Regulation of fungal secondary metabolism. *Nat Rev Microbiol* **11**, 21-32, (2013).

4. Fischbach, M. A. & Walsh, C. T. Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chem Rev* **106**, 3468-3496, (2006).

5. Jenke-Kodama, H. & Dittmann, E. Bioinformatic perspectives on NRPS/PKS megasynthases: advances and challenges. *Nat Prod Rep* **26**, 874-883, (2009).

6. Walsh, C. T., O'Brien, R. V. & Khosla, C. Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds. *Angew Chem Int Ed Engl* **52**, 7098-7124, (2013).

7. Forseth, R. R. *et al.* Homologous NRPS-like Gene Clusters Mediate Redundant Small-Molecule Biosynthesis in Aspergillus flavus. *Angew Chem Int Ed Engl* **52**, 1590-1594, (2013).

8. Wang, M., Beissner, M. & Zhao, H. Aryl-aldehyde formation in fungal polyketides: Discovery and characterization of a distinct biosynthetic mechanism. *Chem Biol* **21**, 257-263, (2014).

9. Hagel, J. M. & Facchini, P. J. Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world. *Plant Cell Physiol* **54**, 647-672, (2013).

10. Bentley, K. W. β-phenylethylamines and the isoquinoline alkaloids. *Nat Prod Rep* **23**, 444-463, (2006).

11. Bok, J. W. *et al.* LaeA, a regulator of morphogenetic fungal virulence factors. *Eukaryot Cell* **4**, 1574-1582, (2005).

12. Bok, J. W. & Keller, N. P. LaeA, a regulator of secondary metabolism in Aspergillus spp. *Eukaryot Cell* **3**, 527-535 (2004).

13. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Research* **43**, 222-226, (2015).

14. MacPherson, S., Larochelle, M. & Turcotte, B. A fungal family of transcriptional regulators: the zinc cluster proteins. *Microbiol Mol Biol Rev* **70**, 583-604, (2006).

15. Forseth, R. R. & Schroeder, F. C. NMR-spectroscopic analysis of mixtures: from structure to function. *Curr Opin Chem Biol* **15**, 38-47, (2011).

16. Forseth, R. R. & Schroeder, F. C. Correlating secondary metabolite production with genetic changes using differential analysis of 2D NMR spectra. *Methods Mol Biol* **944**, 207-219, (2012).

17. Kutchan, T. M. & Dittrich, H. Characterization and Mechanism of the Berberine Bridge Enzyme, a Covalently Flavinylated Oxidase of Benzophenanthridine Alkaloid Biosynthesis in Plants. *J Biol Chem* **270**, 24475-24481, (1995).

18. Winkler, A. *et al.* A concerted mechanism for berberine bridge enzyme. *Nat Chem Biol* **4**, 739-741, (2008).

19. Gaweska, H. M., Roberts, K. M. & Fitzpatrick, P. F. Isotope effects suggest a stepwise mechanism for Berberine Bridge Enzyme. *Biochemistry* **51**, 7342-7347, (2012).

20. Scrutton, N. S. Chemical aspects of amine oxidation by flavoprotein enzymes. *Nat Prod Rep* **21**, 722-730, (2004).

21. Dunn, R. V., Munro, A. W., Turner, N. J., Rigby, S. E. & Scrutton, N. S. Tyrosyl radical formation and propagation in flavin dependent monoamine oxidases. *Chembiochem* **11**, 1228-1231, (2010).

22. Leonard, E., Runguphan, W., O'Connor, S. & Prather, K. J. Opportunities in metabolic engineering to facilitate scalable alkaloid production. *Nat Chem Biol* **5**, 292-300, (2009).

23. Glenn, W. S., Runguphan, W. & O'Connor, S. E. Recent progress in the metabolic engineering of alkaloids in plant systems. *Curr Opin Biotechnol* **24**, 354-365, (2013).

24. Nakagawa, A. *et al.* A bacterial platform for fermentative production of plant alkaloids. *Nat Commun* **2**, 326, (2011).

25. DeLoache, W. C. *et al.* An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. *Nat Chem Biol* **11**, 465-471, (2015).

26. Hawkins, K. M. & Smolke, C. D. Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae. *Nat Chem Biol* **4**, 564-573, (2008).

27. Fossati, E. *et al.* Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in Saccharomyces cerevisiae. *Nat Commun* **5**, 3283, (2014).

28. Thodey, K., Galanie, S. & Smolke, C. D. A microbial biomanufacturing platform for natural and semisynthetic opioids. *Nat Chem Biol* **10**, 837-844, (2014).

29. Galanie, S. & Smolke, C. D. Optimization of yeast-based production of medicinal protoberberine alkaloids. *Microb Cell Fact* **14**, 144, (2015).

30. Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* **349**, 1095-1100, (2015).

31. Imamura, K. *et al.* Identification of a Gene Involved in the Synthesis of a Dipeptidyl Peptidase IV Inhibitor in Aspergillus oryzae. *Appl Environ Microbiol* **78**, 6996-7002, (2012).

32. Macheleidt, J. *et al.* Transcriptome analysis of cyclic AMP-dependent protein kinase A–regulated genes reveals the production of the novel natural compound fumipyrrole by Aspergillus fumigatus. *Mol Microbiol* **96**, 148-162, (2015).

33. Yamada, T. *et al.* Crystal Structure of Serine Dehydratase from Rat Liver. *Biochemistry* **42**, 12854-12865, (2003).

34. Wallner, S. *et al.* Catalytic and Structural Role of a Conserved Active Site Histidine in Berberine Bridge Enzyme. *Biochemistry* **51**, 6139-6147, (2012).

35. Mukherjee, A., Walker, J., Weyant, K. B. & Schroeder, C. M. Characterization of flavin-based fluorescent proteins: an emerging class of fluorescent reporters. *PLoS One* **8**, e64753, (2013).

36. Schneider, P., Weber, M., Rosenberger, K. & Hoffmeister, D. A One-Pot Chemoenzymatic Synthesis for the Universal Precursor of Antidiabetes and Antiviral Bis-Indolylquinones. *Chem Biol* **14**, 635-644 (2007).

37. Fawaz, M. V., Topper, M. & Firestine, S. M. The ATP-Grasp Enzymes. *Bioorg Chem* **39**, 185-191, (2011).

38. Fan, C., Moews, P. C., Shi, Y., Walsh, C. T. & Knox, J. R. A common fold for peptide synthetases cleaving ATP to ADP: glutathione synthetase and D-alanine:d-alanine ligase of Escherichia coli. *Proc Natl Acad Sci U S A* **92**, 1172-1176 (1995).

39. Noike, M. *et al.* A peptide ligase and the ribosome cooperate to synthesize the peptide pheganomycin. *Nat Chem Biol* **11**, 71-76, (2015).

40. Tobiasen, C. *et al.* Nonribosomal peptide synthetase (NPS) genes in Fusarium graminearum, F. culmorum and F. pseudograminearium and identification of NPS2 as the producer of ferricrocin. *Current Genetics* **51**, 43-58, (2007).

41. Nagano, N. *et al.* Class of cyclic ribosomal peptide synthetic genes in filamentous fungi. *Fungal Genet Biol* **86**, 58-70 (2016).

42. Wiemann, P. *et al*. Perturbations in small molecule synthesis uncovers an iron-responsive secondary metabolite network in Apsergillus fumigatus. *Front Microbiol* **5**, 530, (2014).

CHAPTER 4

BIOSYNTHESIS OF THE STRUCTURALLY UNPRECEDENTED IMIZOQUINS IN THE PLANT PATHOGENIC FUNGUS ASPERGILLUS FLAVUS

Abstract: Identification of fungal isoquinoline alkaloids has remained scarce despite the vast number of these compounds that have been characterized from plants. It was recently shown that the human pathogen *Aspergillus fumigatus* biosynthesizes a unique class of fungal isoquinolines, called the Fumisoquins. Formation of the isoquinoline ring present in the structure of the Fumisoquins was shown to occur via a three-step enzymatic mechanism directly analogous to the mechanism used by plants. While the genes responsible for fungal isoquinoline formation lack any homology with the biosynthetic genes in plants, they appear to be conserved across diverse fungi. Here we selected one such predicted isoquinoline producing cryptic biosynthetic gene cluster (BGC) from the plant pathogenic fungus *A. flavus* for further characterization, which we call the *imq* cluster. LC-MS comparative metabolomics followed by detailed 2D NMR spectroscopic analysis of overexpression strains of the transcription factor ImqK revealed *imq*-dependent production of a class of tripeptide-derived alkaloids, the imizoquins. These compound include the expected isoquinoline moiety as part of a tricyclic imidazo[2,1-*a*]isoquinoline ring system, for which there are no other examples among known natural products. As in the case of the fumisoquins, the isoquinoline portion of the imizoquins appears to be derived from hydroxylation, *N*-methylation and subsequent cyclization of tyrosine. This work corroborates the prediction that isoquinoline biosynthesis is conserved

across diverse fungi and highlights the plethora of remaining BGCs that likely produce novel compounds that are assembled around the isoquinoline scaffold.

Introduction: *Aspergillus flavus* is one of the most notorious and deleterious plant pathogens in the world[1]. One mechanism by which *A. flavus* exhibits its high-level of pathogenicity towards plants is through the production of small molecules (SMs) by secondary metabolic biosynthetic gene clusters (BGCs)[2]. The most well-known SM produced by *A. flavus* is Aflatoxin, a carcinogenic polyketide, which, conservatively, is responsible for millions of dollars in crop loss each year in the United States alone[3,4]. Aflatoxin production and expression of its BGC, is positively affected by the global secondary metabolic regulator and virulence factor LaeA[5,6]. In addition to the Aflatoxin BGC, LaeA also positively affects the expression of other, yet to be characterized BGCs. Here we sought to characterize one such LaeA regulated cryptic BGC from *A. flavus*.

A. flavus BGC regulated by LaeA produces isoquinolines: In our search for other LaeA regulated and putative isoquinoline producing gene clusters we identified an eleven-gene cluster, which we named the *imq* cluster (**Figure 4.1a**). The eleven cluster genes encode two small proteins with no predicted function (*imqA* and *imqI*) a canonical three-module non-ribosomal peptide synthetase (NRPS, *imqB*), a phenol 2-monooxygenase (*imqC*), two transporters (*imqD* and *imqJ*), a gibberilin oxidase (*imqE*), an *N*-methyltransferase and an *O*-methyltransferase (*imqF* and *imqG*, respectively), a FAD dependent oxidase (*imqH*), and a C6 transcription factor (*imqK*) (**Figure 4.1a**). Notably, *imqC*, *imqF*, and *imqH* are homologs of *fsqB*, *fsqC*, and *fsqG*, which carry out the three enzyme transformation of tyrosine into tetrahydroisoquinoline *en route* to the Fumisoquins[7]. In order to facilitate chemical characterization of the *imq*-associated metabolites we choose to make an *A. flavus* overexpression (OE) strain for the putative

C6 transcription factor ImqK. Often SM producing BGCs are regulated by C6 transcription factors, enabling whole cluster up-regulation by OE of that single gene[8]. Replacement of the *imqK* promoter with the constitutive *gdpA* promoter (OE::*imqK*) resulted in induction of the other 10 cluster genes (**Figure 4.1b**) and yielded a purple pigmented strain (**Figure D1**). Two cluster flanking genes, (AFLA_064220 and AFLA_064340) were not regulated

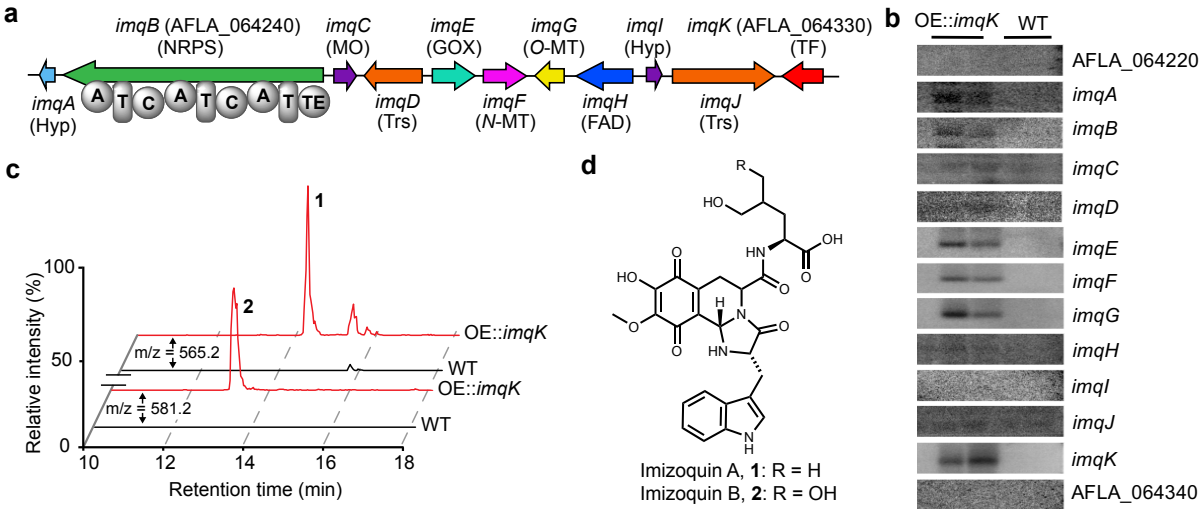


Figure 4.1: **ImqK controls *imq* cluster gene expression and metabolite production.** (**a**) imq gene cluster and putative assignments: Hyp: hypothetical protein; NRPS: non-ribosomal peptide synthetase (A = adenylation, T = thiolation, C = condensation, TE = thioesterase domain): MO: phenol 2-monooxygenase; Trs: transporter; GOX: giberillin oxidase; *N*-MT: *N*-methyltransferase; *O*-MT: *O*-methyltransferase; FAD: FAD-binding domain protein; TF: transcription factor. (**b**) northern blot analysis of *imq* cluster and flanking genes in OE::*imqK* and WT strains. (**c**) LC-MS extracted-ion-chromatograms for **1-2** from analysis of crude extracts from OE::*imqK* and WT. (**d**) structure of Imizoquins A and B, **1-2**, respectively.

by ImqK and thus likely define the boundaries of the cluster (**Figure 4.1b**).

Imizoquin A and B are the major products of the *imq* BGC: To probe the biosynthetic function of the *imq* cluster we employed liquid chromatography – mass spectrometry (LC-MS)-based comparative metabolomics[9]. We compared whole metabolome extracts of the OE::*imqK* strain to WT *A. flavus*. Since the *imq* cluster is not expressed in WT under

laboratory conditions we expected to observe complete differential production of *imq*-associated metabolites in the OE::*imqK* strain relative to WT. Given that liquid cultures of the OE::*imqK* strain had a purple pigmentation, we utilized the visible-light detector during our LC-MS analysis to search for peaks with an absorbance corresponding to a purple pigment (420 nm). By this method we found two starkly differential components of the whole metabolome of OE::*imqK* compared to WT, with low-resolution mass-to-charge ratios of 565.2 and 581.2 (Imizoquin A, **1** and B, **2**, **Figure 4.1c-d**). To determine the structure of **1-2** we grew large scale liquid cultures of OE::*imqK* and conducted several rounds of chromatographic enrichment (see **Appendix D** for details). Following chromatographic enrichment and a high-resolution mass spectrometric analysis (**Table D1**), a standard suite of 2D-NMR spectroscopic experiments was utilized to obtain full structural information for **1-2**.

The imizoquins are tri-peptide imidazo[2,1-*a*]isoquinolines: Analysis of the ^{1}H and dqfCOSY spectra for **1-2** revealed very similar spectroscopic features for both compounds, as expected because their high-resolution mass-to-charge ratios differ by only a single oxygen (**Table D1** and **Appendix D Spectra**). Therefore, we will limit the description of our structural elucidation to **1** (**Figure 4.2**). The aromatic region of the ^{1}H and dqfCOSY spectra for **1** showed two doublets, H-19 and H-22 (δ_H 7.50 and 7.12, respectively) that each coupled to a triplet, H-20 and H-21 (δ_H 6.89 and 6.99, respectively), which also coupled with each other. This in addition to the presence of a singlet, H-16 (δ_H 6.93) was strongly indicative of a substituted indole moiety, such as a tryptophan. H-16 showed a long-range ^{1}H-^{1}H coupling to a diastereostopic methylene, H-

14a/b (δ_H 3.05 and 3.12), which had a 13.8 Hz germinal coupling constant and one additional ^{1}H-^{1}H coupling to a methine, H-6 (δ_H 4.20), entirely consistent with the beta position of an amino acid. HMBC correlations from H-16 to C-14 and the other indole carbons C-18 and C-23 unequivocally confirmed the presence of a tryptophan moiety in the structure of **1**.

In order to extend out from the tryptophan spin-system into the rest of the molecule we looked at long-range ^{1}H-^{1}H couplings of the alpha proton (H-6) of the tryptophan moiety. In addition to coupling with H-14, H-6 also showed a weak ^{1}H-^{1}H coupling with a methine, H-8 (δ_H 5.53) and showed an HMBC correlation to a carbonyl with a chemical shift consistent of cyclic amides, C-5 (δ_C 176.44), suggesting that H-6 and H-8 were part

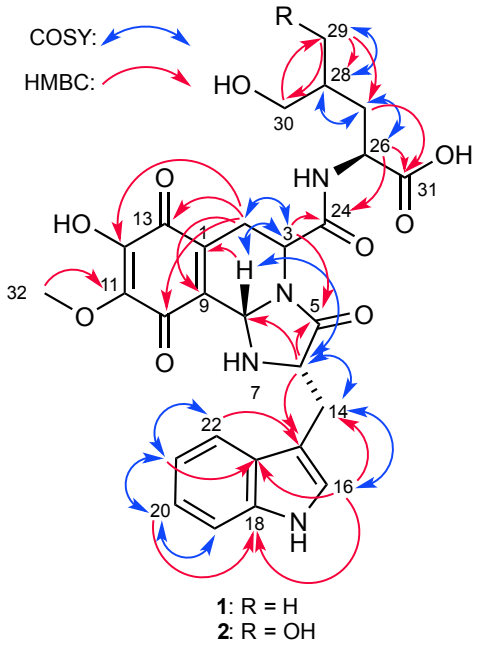


Figure 4.2: **Structure elucidation of 1 and 2.** Showing key COSY correlations (blue double-headed arrows) and key HMBC correlations from each proton to the indicated carbon (red single-headed arrows).

of a fused imidazolidine ring system. H-8 showed an addition long-range ^{1}H-^{1}H coupling to another methine, H-3 (δ_H 4.90), which showed a strong ^{1}H-^{1}H coupling to a

diastereotopic methylene, H-2a/b (δ_H 2.67 and 1.54). H-2a and H-2b had a germinal coupling constant of 17.9 Hz, suggesting they were part of a bi-cyclic fused heteroaromatic ring system, such as tetrahydroisoquinoline[7]. HMBC correlations from H-2a/b showed two carbonyls, C-10 (δ_C 183.50) and C-13 (δ_C 184.22), as well as two quaternary carbons, C-1 (δ_C 134.65) and C-9 (δ_C 145.92). H-8 also showed an HMBC correlation to C-1 and the broadness of the signal for H-8 meant this was likely a three-bond HMBC correlation (three-bond HMBC correlations are strongest in cyclic spin-systems), which placed C-1 three bonds away from H-8. This in addition to the presence of a methoxy group, H_3-32 (δ_H 3.86), which showed an HMBC correlation to a quaternary carbon, C-12 (δ_C 139.23) indicated the structure of **1** contained a highly oxidized *O*-methoxy quinone. The HMBC correlation from H-3 to C-5 confirmed the presence of a imidazo[2,1-*a*]isoquinoline ring system in **1-2**.

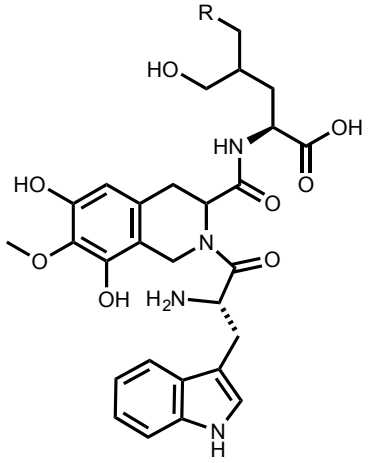
Having established the tri-cyclic imidazo[2,1-*a*]isoquinoline ring system joined to tryptophan we looked for connectivity to the last remaining portion of the molecule. The HR-MS derived molecular formula, $C_{28}H_{30}N_4O_9$, suggested we were missing a structural feature with a molecular formula of $C_6H_{13}NO_3$. The missing formula corresponded nicely to hydroxy-leucine (in the case of **1**) or dihydroxy-leucine (for **2**). This in addition to the peptidic nature of **1-2** suggested that the last point of connection would likely be the carboxy group in the three-position of the isoquinoline ring (derived from L-tyrosine incorporation). Indeed, another methine, H-26 (δ_H 4.28) showed an HMBC correlation to the tyrosine derived carbonyl, C-24 (δ_C 171.24). H-26 had a ^{1}H-^{1}H coupling with a methylene, H-27a/b (δ_H 1.57 and 1.77), which had only one other ^{1}H-^{1}H coupling to methine, H-28 (δ_H 1.57). H-26 and H-27a/b both showed an HMBC correlation to the last

carbonyl present in the structure of **1-2**, C-31 (δ_C 175.92), which was not observed by any other protons, suggesting it was the free carboxylic acid. This was further confirmed by the propensity of **1-2** to ionize in electrospray negative ionization $[M-H]^-$. In addition to ^1H-^1H coupling with H-27a/b, H-28 also showed a ^1H-^1H coupling to a terminal methylene, H_2-30 (δ_H 3.31) and a methyl, H_3-29 (δ_H 0.84), which collectively, confirmed the incorporation of a 5-hydroxy-L-leucine in the structure of **1** and 5,5-dihydroxy-L-leucine in the structure of **2** (see **Table D4-5** for complete chemical shift table).

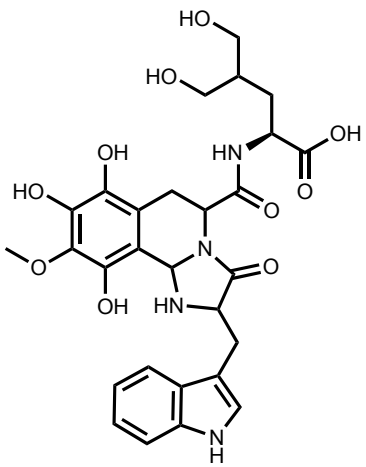
Imizoquins A and B are structurally unprecedented natural products: Upon completion of the structural elucidation of **1-2** we searched Reaxys and Scifinder databases to find related compounds derived from natural sources. Our search revealed a structurally-related set of isoquinoline alkaloids derived from *A. oryzae* called TMC-2A-B (**3-4**, **Figure 4.3**), which were identified in 1997 by Asai and coworkers[10]. TMC-2A-B were shown to be dipeptidyl peptidase IV inhibitors. Additionally, our search showed that more recently, in 2012, Imamura and coworkers identified a gene involved in the biosynthesis of **3-4**, which at that time they renamed **3** as Wyk-1[11]. The gene identified by Imamura and coworkers is a part of a BGC in *A. oryzae* that also possesses homologs to *fsqB*, *fsqC*, and *fsqG,* further confirming the conserved nature of isoquinoline production across a range of fungal species[7]. The realization of the existence of **3-4** prompted us to revisit our LC-HRMS analysis of whole metabolome extracts from liquid cultures of the OE::*imqK* strain. Somewhat unsurprisingly, we found that the OE::*imqK* strain also produces appreciable quantities of **3-4** (**Figure D2**), suggesting that **3-4** are intermediates or shunt metabolites in the *imq* pathway, as they lack the imidazolidine ring

system present in **1-2**. The structural distinction between **1-2** and **3-4** is important because the simple isoquinoline motif in **3-4** is widespread among natural products, however, the fused tricyclic imidazo[2,1-*a*]isoquinoline ring system in **1-2**, to the best of our knowledge does not exist among other known natural products.

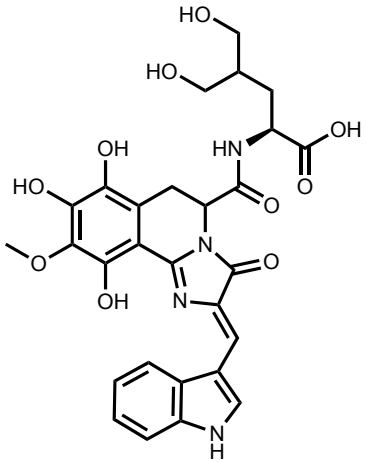
Imizoquin A and B could serve as ROS quenching metabolites in fungi: When we were conducting our LC-HRMS analysis to survey the OE::*imqK* strain for the previously identified isoquinolines (**3-4**) we also identified other features in the metabolome that appeared to be *imq*-associated metabolites based on their high-resolution mass-to-charge ratios of 583.2046 and 579.1733. Chromatographic enrichment (**Appendix D**) of these new *imq*-associated features followed by full 2D-NMR spectroscopic



TMC-2A (Wyk-1), **3**: R = OH
TMC-2B, **4**: R = H



Imizoquin C, **5**

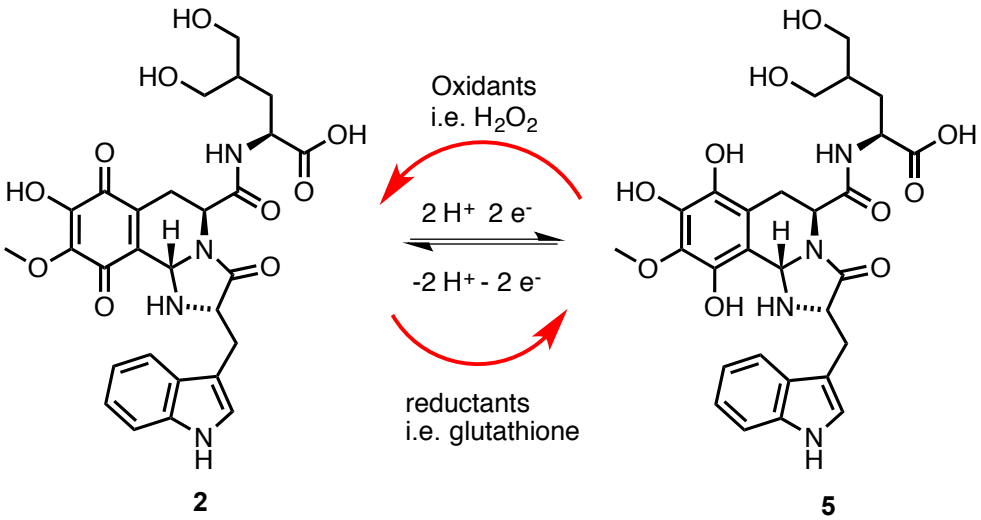


Imizoquin D, **6**

Figure 4.3: Structures of previously identified isoquinolines (3-4) and novel Imizoquins (4-5). Stereochemistry is not defined for the imidazo[2,1-a]isoquinoline ring system of Imizoquin C because we currently lack spectroscopic evidence to support a specific configuration.

characterization revealed they were indeed related to **1-2**, and each represented a different oxidation state of the same scaffold (**Table D6-7** for NMR assignments and

Appendix D for NMR spectra). In accordance with standard naming routines we called these compounds Imizoquin C and Imizoquin D (**5** and **6**, respectively, **Figure 4.3**). The structural difference between **2** and **5** is a quinone to hydroquinone (**2** to **5**), suggesting that **1-2** may serve as a role in fungal redox chemistry (**Figure 4.4**). To test this hypothesis we isolated several milligrams of **1** from OE::*imqK* (see **Appendix D** for details) cultures and subjected the purified compound to a total antioxidant assay (TAC, **Figure D3**)[12]. In the TAC assay **1** showed a concentration dependent increase in antioxidant capacity (**Figure D3**). The amount of TAC exhibited by **1** was comparable to uric acid (a standard measure of TAC) for concentrations as low as 3 μM, which is consistent with isolated quantities of Imizoquins A and B (2-4 mg / L of liquid culture). In addition to Imizoquin A and B it is likely the full suite of *imq*-dependent metabolites also have antioxidant capacity, suggesting a possible biological function for the *imq* cluster is to produce ROS quenching metabolites[13]. This could be in an effort to provide resistance to the solar radiation-induced ROS species that *A. flavus* experiences during growth on plant surfaces[14]. We



2 **5**

Figure 4.4: Mechanism showing how imizoquins could serve a role in fungal redox: Conversion of **2** into **5** through the reduction of the quinone to the hydroquinone, for example by glutathione and vice-versa oxidation **5** into **2** through ROS species such as H_2O_2, converting the hydroquinone back to the quinone.

hope to explore the biological activity of the imizoquins more in the future, however, the existing data supports the increasingly realized essential role of secondary metabolite BGC's in fungal biology[15].

***Imq*-dependent metabolite profile reveals a model for imizoquin biogenesis:** Once we had established the major and minor products of the *imq* custer and a potential biological role for these metabolites we sought to understand their biogenesis. Since, part of the *imq* cluster shares high-homology with tailoring genes in the *fsq* cluster, the mechanism for isoquinoline formation *en route* to the imizoquins did not need further exploration. Normally, to gain insights into the role of each individual gene (and therefore the enzyme encoded) in a BGC we would construct double-mutants with individual genes deleted in the OE::*imqK* background[16]. This however, seemed like an extensive task for the *imq* pathway, since the homology for many of the tailoring enzymes offered enough information to provide an accurate putative function in the biogenesis of **1-2**. Furthermore,

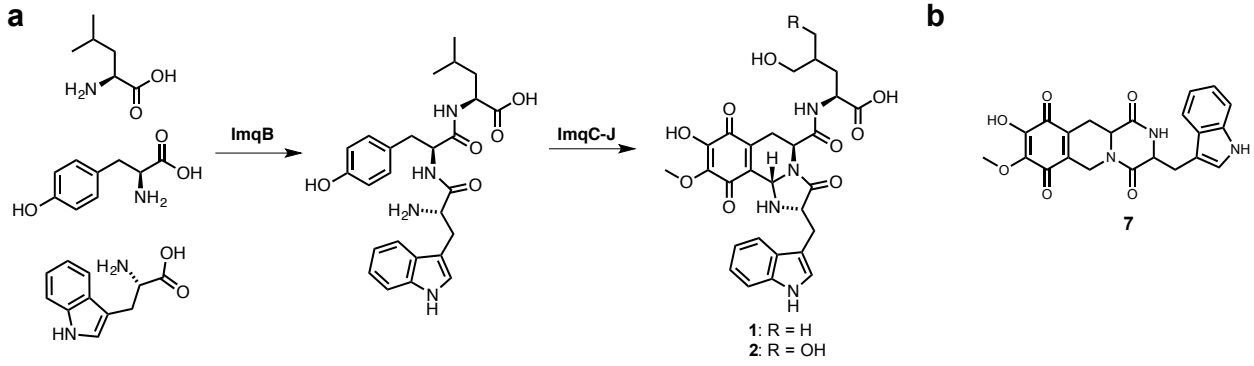


Figure 4.5: Potential model for biogenesis of 1-2 and structure of shunt metabolite 7. Stereochemistry is not defined for 7 because we currently lack spectroscopic evidence to support a specific configuration.

unlike the *fsq* pathway (Chapter 3) the NRPS, *imqB*, is a canonical three-module NRPS with a terminal thioesterase domain (TE, **Figure 4.1a**). This indicated that the assembly of the tripeptide backbone of the imizoquins likely occurs in a very predictable manner

where C1 and C2 catalyze the formation of two peptide bonds, first between T1-tethered L-Trp and T2-tethered L-Tyr, then T2-tethered L-Trp-L-Tyr dipeptide is captured by L-Leu on T3. Lastly, the T3-tethered tripeptide (L-Trp-L-Tyr-L-Leu) is offloaded from the NRPS by the TE domain and subsequently modified by *imq* tailoring enzymes (**Figure 4.5a**). If this model was accurate, the substrate for each of the *imq* tailoring enzymes would be the tripeptide, NH2-L-Trp-L-Tyr-L-Leu-OH. Evidence to the contrary of this simple prediction for *imq* biogenesis came upon the discovery of another *imq*-associated metabolite (**7**, **Figure 4.5b** and see **Appendix D** for HR-MS and NMR assignments**)**

There are a few possibilities that come to mind for an *imq* biosynthetic pathway that is consistent with the formation of **7**. One, an early termination event where the amino group of L-Trp cyclizes intramolecularly and releases cyclo(L-Tyr-L-Trp) from ImqB, which

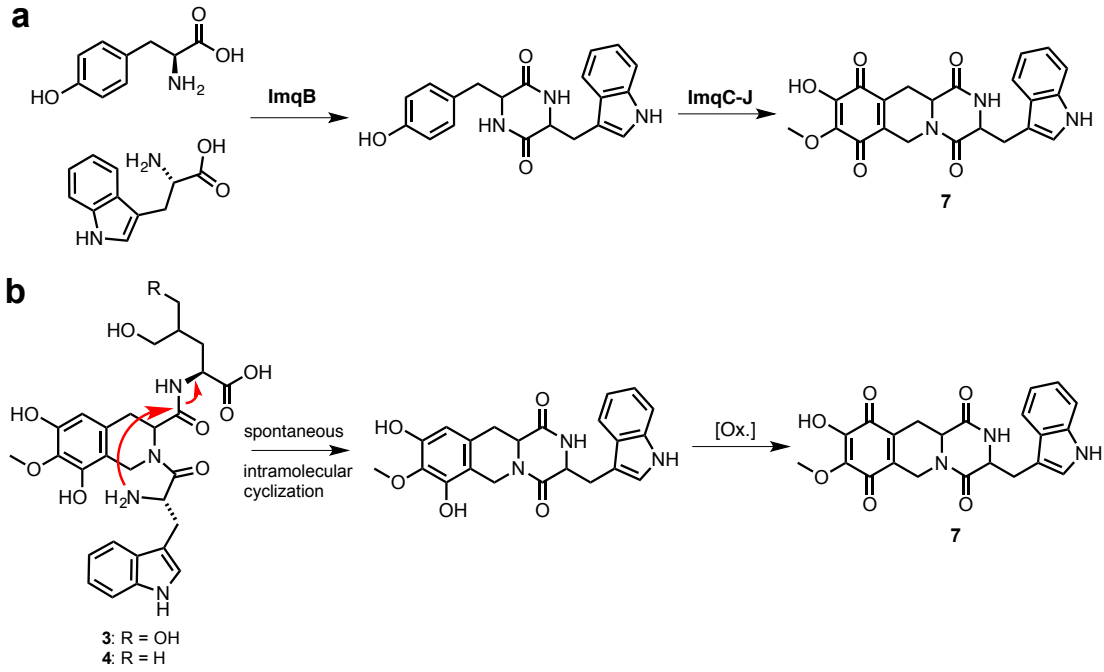


Figure 4.6: Possible mechanisms explaining the existence of 7 in OE::*imqK* strains. (**a**) L-Tyr and L-Trp spontaneously release from T2 of ImqB forming cyclo(L-Tyr-L-Trp), which is then modified by *imq* cluster tailoring enzymes to form **7**. (**b**) Compound 3-4 are produced as speculated in **Figure 4.5**, then upon standing undergo spontaneous cyclization and oxidation ([Ox.]) to form **7**.

is then acted upon by the *imq* cluster tailoring enzymes (**Figure 4.6a**). Two, the L-Trp-derived amino group in **3-4** cyclizes intramolecularly and releases hydroxy-L-Leu or dihydroxy-L-Leu, then further oxidation of the isoquinoline ring results in formation of **7** (**Figure 4.6b**). This possibility seems the least likely, since we didn't observe formation of **7** in isolated samples of **3-4** (spectra, **Appendix D**) and cleavage of the L-Leu amide bond is unfavorable. Lastly, and perhaps most likely, the modifications to form the oxidized isoquinoline portion of **1-2** occur while the dipeptide is tethered to ImqB, which occasionally prematurely cyclizes off the T2 domain of ImqB, resulting in appreciable quantities of **7 (Figure 4.7)**. This is a preferred model over the one shown in **Figure 4.6a** because it doesn't require extreme substrate promiscuity for the *imq* tailoring enzymes. If cyclo(L-Tyr-L-Trp) was modified in addition to linear L-Trp-L-Tyr-L-Leu, as would be suggested by the model in **Figure 4.6a**, then each tailoring enzyme would need to accept two different substrates with dramatically different size and shape. Collectively, our characterization of the *fsq* pathway (where the tailoring modification occur on a tethered dipeptide) in addition to the observation of **7** lead to an *fsq*-analogous model for the *imq* pathway where **1-2** biogenesis occurs via isoquinoline formation on a T2-tethered L-Trp-

L-Tyr prior to the formation of the tripeptide (**Figure 4.7**). Following action of ImqC, ImqF, and ImqG, the isoquinoline dipeptide is moved from the T2 domain of ImqB and T3-tethered L-Leu is condensed in via the C2 domain, resulting in the near complete **1-2** (**Figure 4.7**). Lastly, ImqD, the gibberilin oxidase, either partially (resulting in **2**) or completely hydroxylates the four position of the L-Leu moiety, resulting in **1** (**Figure 4.7**).

Conclusions: Our characterization of the *imq* BGC in *A. flavus* represents another confirmed example of conserved isoquinoline production in diverse fungi. The biological function of the imizoquins remains in question, however, our characterization of several different *imq* dependent metabolites with varied extent of oxidation indicates a role in

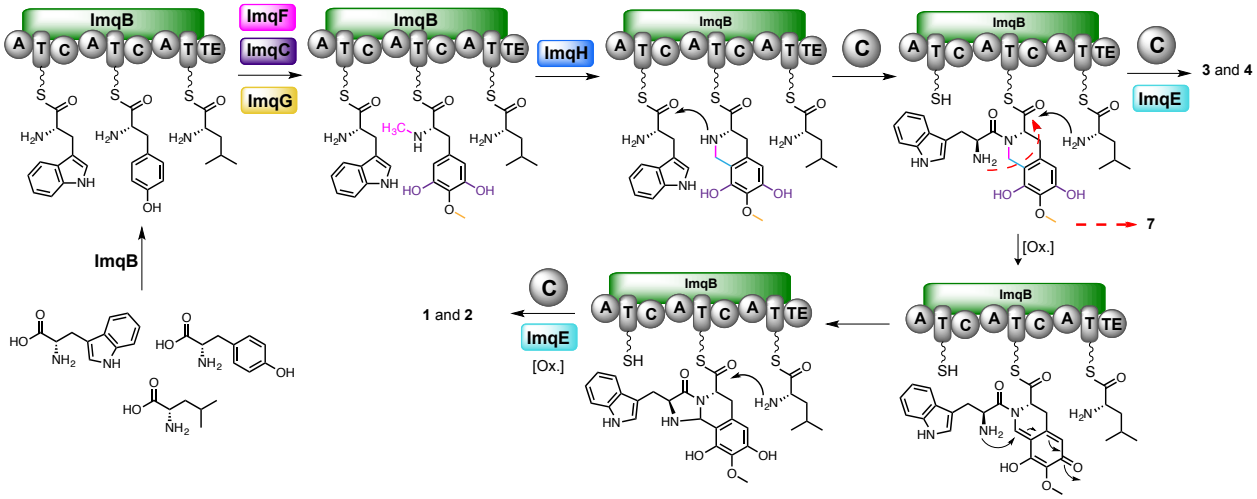


Figure 4.7: Model for *imq* biosynthesis. L-tryptophan, L-tyrosine, and L-leucine are loaded onto ImqB thiolation (T) domains via their respective adenylation (A) domains. Then ImqF (N-methyltransferse, pink), ImqC (phenol 2-monooxygenase, purple), and ImqG (O-methyltransferase, yellow) act on the T2-bound L-Tyr. Following N-methylation and hydroxylation ImqH (FAD dependent oxidase, blue) catalyzes the formation of the isoquinoline ring. Then C2 condenses L-Leu into the modified T2-tethered dipeptide, which subsequently is hydroxylated by ImqE (giberillin oxidase, turquoise) to form **3** and **4**. Alternatively, the modified T2-tethered dipeptide can prematurely cyclize off (red arrow) to form **7**. Lastly, prior to L-leu condensation another oxidation precedes the L-Leu capture, allowing the formation of **1** and **2**.

redox chemistry. Further supporting this hypothesis is the total antioxidant capacity for the purified compounds. It should be noted that the identification of many of the oxidation-sensitive *imq* cluster metabolites was greatly facilitated by the use of 2D NMR-based metabolomics, highlighting the utility of this method for comprehensive analysis of new biosynthetic pathways.

REFERENCES

1. KLICH, M. A. Aspergillus flavus: the major producer of aflatoxin. *Mol Plant Pathol* **8,** 713–722, (2007).

2. Amaike, S. & Keller, N. P. Aspergillus flavus. *Annu Rev Phytopathol* **49,** 107–133, (2011).

3. Bhatnagar, D., Ehrlich, K. C. & Cleveland, T. E. Molecular genetic analysis and regulation of aflatoxin biosynthesis. *Appl Microbiol Biotechnol* **61,** 83–93, (2003).

4. Mitchell, N. J. *et al.* Potential economic losses to the US corn industry from aflatoxin contamination. *Food Additives & Contaminants: Part A* **33,** 540–550, (2016).

5. Amaike, S. & Keller, N. P. Distinct Roles for VeA and LaeA in Development and Pathogenesis of Aspergillus flavus. *Eukaryotic Cell* **8,** 1051–1060, (2009).

6. Bok, J. W. & Keller, N. P. LaeA, a regulator of secondary metabolism in Aspergillus spp. *Eukaryotic Cell* **3,** 527–535, (2004).

7. Baccile, J. A. *et al.* Plant-like biosynthesis of isoquinoline alkaloids in Aspergillus fumigatus. *Nat Chem Biol* **12**, 419-424, (2016).

8. Yin, W. & Keller, N. P. Transcriptional regulatory elements in fungal secondary metabolism. *J Microbiol* **49,** 329–339, (2011).

9. Forseth, R. R. *et al.* Identification of Cryptic Products of the Gliotoxin Gene Cluster Using NMR-Based Comparative Metabolomics and a Model for Gliotoxin Biosynthesis. *J Am Chem Soc* **133,** 9678–9681, (2011).

10. Asai, Y., *et al*. TMC-2A,-2B and-2C, new dipeptidyl peptidase IV inhibitors produced by Aspergillus oryzae A374. II. Isolation and structure determination. *J Antibiot* **8**, 653-658, (1997).

11. Imamura, K. *et al.* Identification of a Gene Involved in the Synthesis of a Dipeptidyl Peptidase IV Inhibitor in Aspergillus oryzae. *Appld Environ Microbiol* **78,** 6996–7002, (2012).

12. Young, I. S. Measurement of total antioxidant capacity. *J Clin Pathol* **5**, 339, (2001).

13. Ames, B. N., Cathcart, R. & Schwiers, E. Uric acid provides an antioxidant defense in humans against oxidant-and radical-caused aging and cancer: a hypothesis. *Proc Natl Acad Sci U S A* **11**, 6858-6862, (1981).

14. Rastogi, R. P., Singh, S. P., Häder, D.-P. & Sinha, R. P. Detection of reactive oxygen species (ROS) by the oxidant-sensing probe 2‘,7’-dichlorodihydrofluorescein diacetate in the cyanobacterium Anabaena variabilis PCC 7937. *Biochem Biophys Res Commun* **397,** 603–607, (2010).

15. Scharf, D. H., Heinekamp, T. & Brakhage, A. A. Human and Plant Fungal Pathogens: The Role of Secondary Metabolites. *PLoS Pathog* **10,** e1003859, (2014).

16. Yin, W. B. *et al.* A Nonribosomal Peptide Synthetase-Derived Iron(III) Complex from the Pathogenic Fungus Aspergillus fumigatus. *J Am Chem Soc* **135,** 2064–2067, (2013).

CHAPTER 5

BIOSYNTHESIS OF THE FUMIZINONES: SECONDARY PRODUCTS OF THE FSQ CLUSTER IN ASPERGILLUS FUMIGATUS

Abstract: Microbial biosynthetic gene clusters (BGCs) containing non-ribosomal peptide synthetases (NRPSs) typically produce a single class of small-molecules (SMs) with limited structural diversity within the context of the cluster products. The outcome of microbial NRPS-based BGCs can therefore be summarized as one or two major products (the most highly-modified SM) and often an array of associated intermediates that follow the same enzymatic entourage. Here we show that the previously characterized *fsq* BGC in *Aspergillus fumigatus* unexpectedly has a secondary set of SM products, the fumizinones, which result from divergent enzymatic activity of the cluster NRPS-like enzyme (FsqF) and ATP-grasp enzyme (FsqD). The fumizinones represent a novel addition to the limited number of pyrazinone natural products identified to date. Stable-isotope incorporation suggests that fumizinone biogenesis proceeds via a unique carbon-carbon bond forming reaction between two L-serine derived precursors. This work represents the first example of a single BGC producing two different structurally unrelated SMs and adds to the mounting evidence that NRPS-like enzymes are a rich and untapped source of new natural products.

Introduction: Historically, small-molecule (SM) production from microbial biosynthetic gene clusters (BGCs) has been, perhaps, the richest source of important natural products[1]. Many of these medicinally important compounds derive from BGCs harboring large canonical non-ribosomal peptide synthetases (NRPSs) or polyketide synthases (PKSs), the megasynth(et)ases[1,2]. In particular, large canonical NRPSs are defined as having multiple modules which are each composed of at a minimum, an adenylation, thiolation, and condensation domain. Frequently, other domains with varied enzymatic activity, such as methyl transferase or oxidase/reductase domains are found within one or more modules[2]. Megasynthetase-based SM biosynthesis proceeds via an assembly line style of loading (adenylation), tethering (thiolation), and elongation (condensation), of amino acids. Termination of the assembly line occurs with a variety of mechanisms (enzymatic and/or non-enzymatic), where either the unmodified, or modified peptide (via tailoring domains or cluster enzymes) is released from the N-terminus of the NRPS[3].

The assembly line nature of canonical NRPSs provides a basis for semi-accurate prediction of the amino acids and modifications that will occur during the biogenesis, making them tempting targets for natural products chemists[4]. Additionally, early discovery efforts from BGC's harboring megasynth(et)ases were rewarded with therapeutic compounds, such as Vancomycin and Erythromycin[5,6]. Accordingly, the vast majority of characterized NRPSs are multi-module versions with peptidyl SM products consisting of several amino acids (or derivatives thereof)[7]. Unfortunately, continued efforts to capture new and important bioactivity from megasynthetase-derived SMs have not been rewarded as expected from the early successes in the field, although exceptions exist, such as Teixobactin[8].

Recently, a related class of enzymes called NRPS-like enzymes have emerged as a source of SMs with overall structural novelty and potentially useful bioactivity. NRPS-like synthetases differ from their larger counter-parts primarily in that they lack complete modules[9,10]. In fact, all NRPS-like enzymes consist of less than one module, which importantly, never contains a condensation domain. The dearth of condensation domains among NRPS-like enzymes is what makes them interesting targets for new molecular assemblies. The absence of a condensation domain significantly hinders one's ability to predict the bonds formed between two amino acids, which ultimately leaves an immense number of possible structural outcomes, even if the substrates are constricted to solely standard amino acids (which may not always be the case)[11]. While relatively little is currently known about the biosynthetic potential of BGCs containing NRPS-like enzymes, evidence from early publications suggest their potential for biosynthesizing complex architectures far exceeds that of their larger counter-parts[9,10]. Moreover, because NRPS-like (and PKS-like) enzymes are small in comparison to the canonical enzymes they have often been overlooked computationally, resulting in a significant underestimation of their true numbers among microbial genomes[12]. Here we sought to highlight the utility of NRPS-like enzymes by characterizing an additional biosynthetic function of the NRPS-like enzyme at the core of the previously characterized *fsq* BGC in the human pathogen *Aspergillus fumigatus*.

The fumizinones are secondary products of the *fsq* cluster: Recently we published the characterization of the Fumisoquins, a novel class of isoquinoline alkaloids derived from the *fsq* BGC[9]. This was accomplished in no small part through the usage of differential analysis by 2D-NMR spectroscopy (DANS) combined with liquid chromatography – high-resolution mass spectrometry (LC-HRMS)-based comparative metabolomics of overexpression (OE) and knock-out (Δ) strains of the cluster transcriptional regulator, *fsqA* (**Figure E1**)[13]. While the Fumisoquins certainly represent the major products of the *fsq* cluster, upon further analysis of our comparison of whole-metabolome extracts of the OE::*fsqA* vs. Δ*fsqA* strains we found two additional features that were consistently produced in an FsqA dependent manner (**1-2**, **Figure 5.1a**). The high-resolution mass-to-charge ratio of these differential features (**1-2**, **Figure 5.1a**) suggested they were not intermediates or shunt metabolites in the previously established *fsq* pathway, since their predicted molecular formulae was consistent with compounds

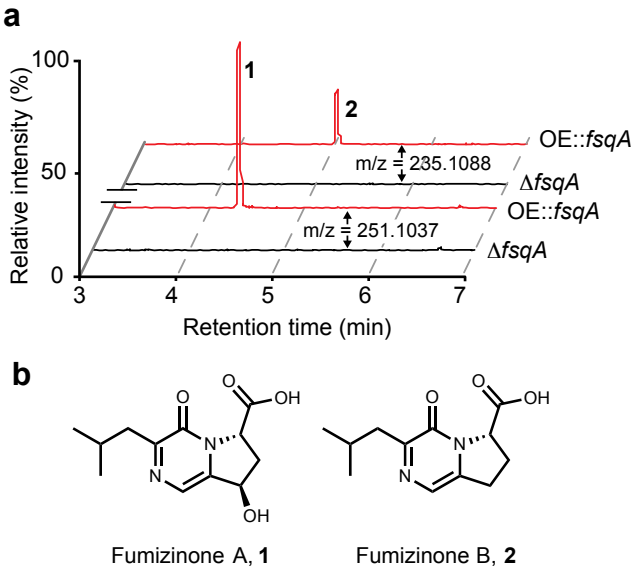


Figure 5.1: FsqA dependent metabolite production and structure of fumizinones. (**a**) LC-HRMS extracted-ion-chromatograms for **1-2** in whole-metabolome extracts of OE::*fsqA* (red lines) and Δ*fsqA* (black lines). (**b**) structure of the **1-2**.

containing two nitrogens, whereas, all Fumisoquins of similar size only posessed one nitrogen. To confirm that **1-2** were in fact structurally unrelated to the other *fsq*-associated metabolites (Fumisoquins and their intermediates, Chapter 3) we grew large scale liquid cultures of OE::*fsqA* and conducted several rounds of chromatographic enrichment (see **Appendix E** for details). Following chromatographic enrichment, a standard suite of 2D-NMR spectroscopic experiments was utilized to obtain full structural information for **1-2**. (**Figure 5.1b**).

The fumizinones are a novel members of the small class of pyrazinone natural products: The peak intensity for **1-2** in our LC-HRMS analysis suggested isolating quantities sufficient for NMR analysis would be rather straightforward. However, after several rounds of isolation from 8 L of liquid culture of the OE::*fsqA* strain we only obtained sub-milligram quantities of **1-2**. Fortunately, we isolated enough **1** to obtain high-quality 2D NMR spectra, affording us the ability to conduct a full spectroscopic characterization (**Figure 5.2**, see **Table E4** for complete chemical shift assignments and **Appendix E** for spectra). Our entry into the spin-system of **1** was from the aliphatic region, where two equivalent methyl groups, H_3-12 and H_3-13, (δ_H 0.93 and 0.93, respectively)

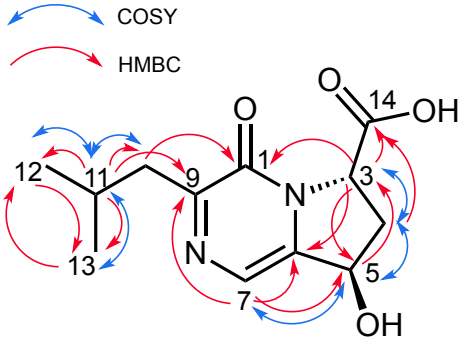


Figure 5.2: Structure elucidation of 1. Showing key COSY correlations (blue double-headed arrows) and key HMBC correlation from each proton to the indicated carbon (red single-headed arrows).

showed a 7 Hz ^{1}H-^{1}H coupling with a methine, H-11 (δ_H 2.20). From H-11 the spin system did not extend much further, as H-11 had only one other coupling to a methylene H-10a/b (δ_H 2.60 and 2.60). The carbon chemical shift of position ten (δ_C 42.37) suggested the presence of an alpha heteroatom or aromatic spin-system. HMBC correlations from H-10a/b to C-9 (δ_C 159.09) and C-1 (δ_C 156.46) confirmed the presence of the suspected adjacent heteroaromatic spin-system.

In order to connect the isopropyl-based spin-system to the remainder of the molecule we next looked at the aromatic region of the ^{1}H NMR spectra. The aromatic region only contained one singlet, H-7 (δ_H 7.46), which showed a strong HMBC correlation to C-9 and a weaker correlation to C-1. H-7 showed one additional HMBC correlation to an aromatic carbon, C-6 (δ_C 144.37), which collectively suggested H-7 was part of a heteroaromatic six-membered ring system, such as a pyrazinone. The dqfCOSY spectrum of **1** revealed that, despite being a singlet in the ^{1}H NMR spectrum, H-7 showed a long-range ^{1}H-^{1}H coupling to a methine H-5 (δ_H 5.15), which appeared to be attached to an oxygen-bearing carbon, C-5 (δ_C 72.43). H-5 showed an HMBC correlation back to C-6 and C-7, solidifying its position three bonds away from H-7. H-5 also showed strong ^{1}H-^{1}H coupling with a diastereotopic methylene H-4a/b (δ_H 2.22 and δ_H 2.72). The germinal ^{1}H coupling constant of H-4a and H-4b (13.5 Hz) suggested that they were part of a five-membered ring. In addition to coupling with H-5, H-4a/b also had a 1H-1H coupling with another methine, H-3 (δ_H 4.83), which showed strong HMBC correlations to two carbonyls, C-1 and C-14 (δ_C 176.02). This connected the two hemispheres of the molecule and confirmed that the structure of **1** was composed of a fused isobutyl-dihydropyrrolo-pyrazinone (**Figure 5.2**)[14].

Stable-isotope amino acid labeling provides insights into fumizinone biogenesis: Once we had established the structure of **1-2**, we sought to understand how they could be biosynthesized in a manner consistent with the Fumisoquins. We started by conducting a stable-isotope amino acid labeling experiment to determine which amino acids are incorporated in **1-2**. First, we tested L-Leucine because that was the most probable candidate for the origin of the isobutyl moiety present in the structure **1-2**. LC-MS analysis of whole-metabolome extracts from cultures of OE::*fsqA* strain fed with L-Leucine and L-[1-^{13}C]-Leucine revealed that, as expected, L-Leucine is incorporated in **1-2** (**Figure E2**). It was rather straightforward to anticipate the incorporation of L-Leucine in **1-2** based on their structures, however, the remainder of the structure does not represent an obvious amino acid. A natural starting point was to test L-Serine and L-Cysteine because we had previously established that the Fumisoquins, which were supposedly derived from the same BGC incorporated L-Serine based substrates into their structures.

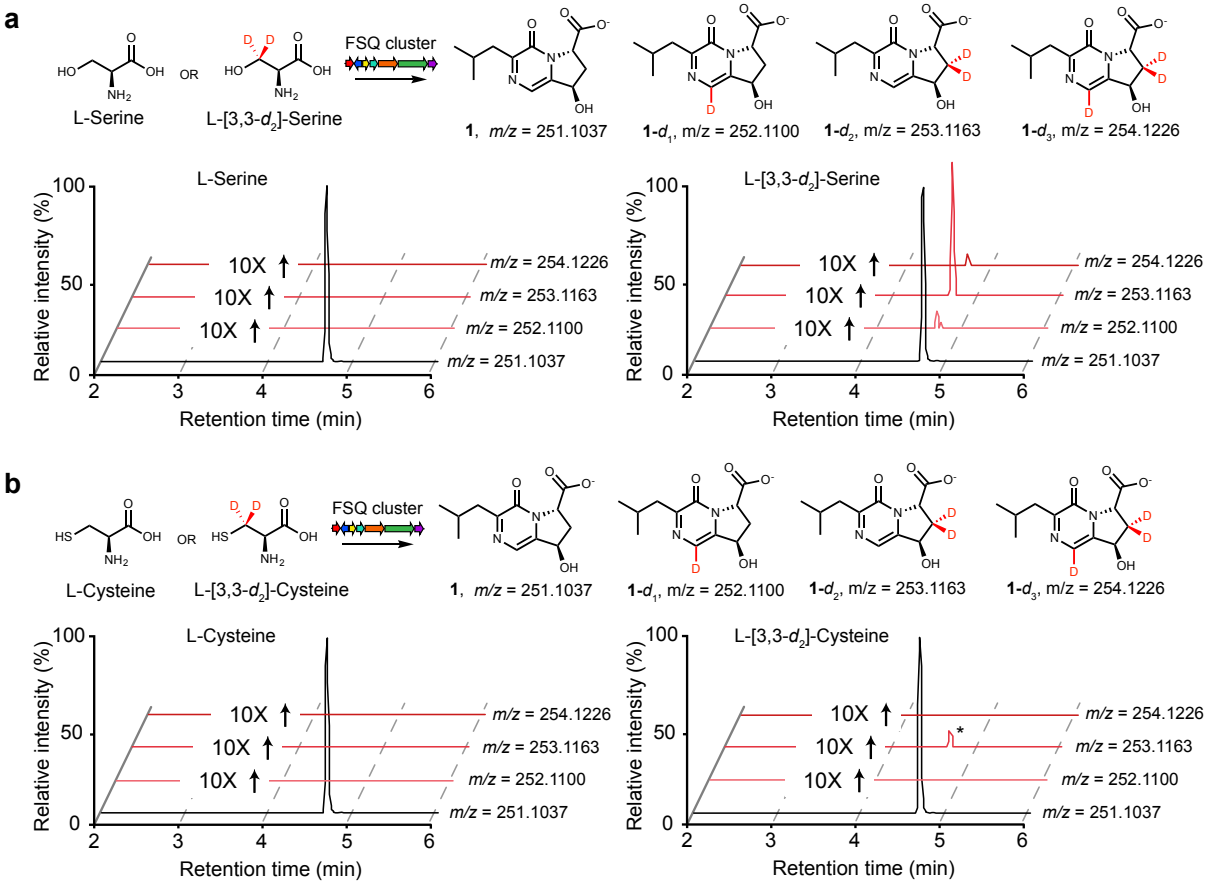


Figure 5.3: Stable-isotope incorporation into 1-2. (**a**) L-Serine or L-[3,3-d_2]-Serine are incorporated into **1-2** with the resulting deuteration shown in red. Below are extracted-ion-chromatograms for the indicated high-resolution mass-to-charge from the control or labelling experiment. (**b**) L-Serine or L-[3,3-d_2]-Serine are incorporated into **1-2** with the resulting deuteration shown in red. Below are extracted-ion-chromatograms for the indicated high-resolution mass-to-charge from the control or labelling experiment.

In order to test for for the incorporation of L-Cysteine and L-Serine we utilized LC-HRMS analysis of OE::*fsqA* cultures fed with L-Cysteine or L-Serine (controls) or L-[3,3-d_2]-Cysteine/Serine (**Figure 5.3a-b**). Interestingly, we observed that not only is one L-Serine incorporated into **1-2** but a second L-Serine is also incorporated (**Figure 5.3a**). Similar to our analysis of the amino acid incorporation in the Fumisoquins we also found that L-Cysteine was not incorporated in **1-2** (**Figure 5.3b**). LC-HRMS analysis was pivotal in distinguishing between deuterated species or natural

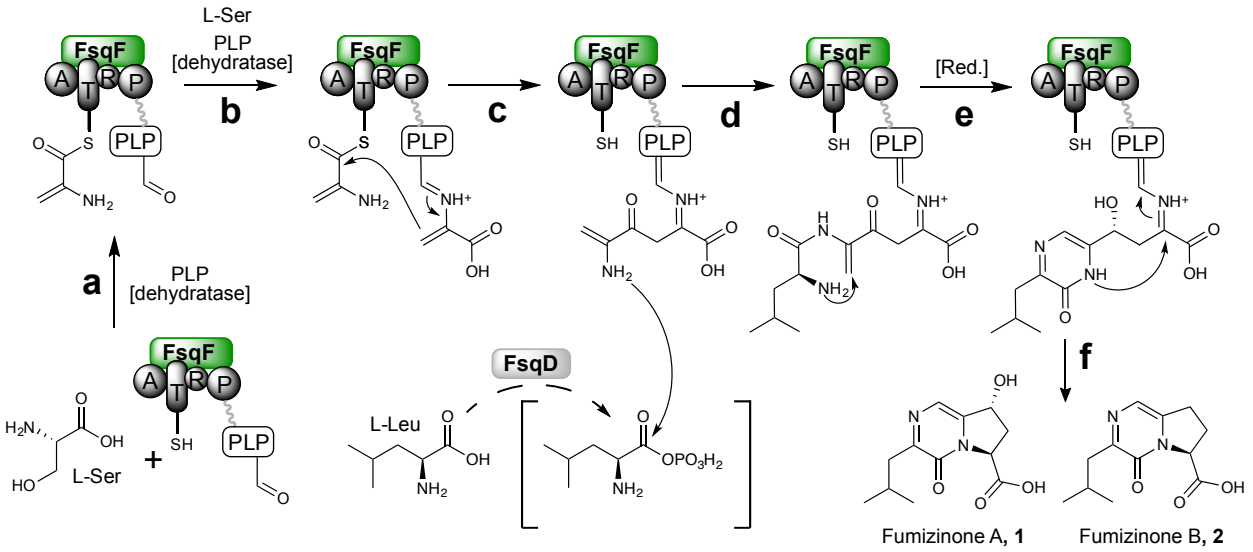


Figure 5.4: Tentative model for fumizinone biogenesis. (**a**) L-Serine is dehydrated by the PLP-dependent domain in FsqF and tethered to the T domain. (**b**) another L-Serine is dehydrated by the PLP-domain and (**c**) subsequently undergoes a claisen condensation with the T-bound dehydroalanine. (**d**) The PLP-bound adduct adds to an activated L-Leucine (perhaps by FsqD), which then (**d**) cyclizes intramolecularly to form the pyrazinone ring. (**e**) one reductive equivalent reduces the ketone to hydroxy and then (**f**) the pyrazinone amide cyclizes, releasing **1-2** from the PLP domain. A = adenylation, T = thiolation, R = reductase and P = PLP-dependent domain.

abundance carbon 13 species, which overlap in a low-resolution mass spectrometry experiment.

The results of the stable-isotope labeling experiments allowed us to propose a tentative model for the biogenesis of **1-2** (**Figure 5.4a-f**). Analogous to the Fumisoquins, one of the early steps in the fumizinone pathway is a carbon-carbon bond forming reaction catalyzed by FsqF. However, between two L-Serine derived dehydroalanine's (DHAs) rather than one dehydroalanine and an activated tyrosine (**Figure 5.4a-b**). Following the FsqF mediated dimerization of dehydroalanine, FsqD provides an activated L-Leucine, which forms a peptide bond with the free amino group of the DHA dimer (**Figure 5.4c-d**). Following peptide-bond formation an intramolecular cyclizes furnishes the pyrazinone moiety (**Figure 5.4e**). Lastly, the L-Serine derived ketone is reduced to the corresponding hydroxy (possibly by action of the reductase domain (R, **Figure 5.4**)) and the pyrazinone

amide bites into the PLP-imine to release **1-2**. This model for fumizinone biogenesis is in contrast to our proposal for the Fumisoquin pathway where FsqD serves to activate L-Tyrosine, which means FsqD exhibits substrate promiscuity to the extent of accepting L-Tyrosine and L-Leucine. One of the only other ATP-grasp enzymes characterized also showed high-substrate promiscuity, giving some credence to this hypothesis[15]. The corresponding fumizinone derivatives where instead of L-Leucine, L-Tyrosine was incorporated were not found in our LC-HRMS survey. Perhaps they were too unstable, or simply don't fit in the active site of FsqF after the dimerization of DHA.

Conclusions: To the best of our knowledge the fumizinones represent the first pyrazinones isolated from *A. fumigatus* and add to the interesting biosynthetic transformations afforded by the under-appreciated class of NRPS-like enzymes. Our observation that two L-Serine's and L-Leucine are incorporated into the fumizinones provided some insight into their biosynthetic pathway, however, work remains to clarify the correct biogenesis. In particular, *in vitro* reconstitution of the enzymes involved, FsqF and FsqD will be critical to determine if they function in a manner consistent with our proposed model. Lastly, the fumizinones are intriguing from the stand-point of a unique BGC where two distinct biosynthetic outcomes are possible. Perhaps regulation of the two divergent pathways is a strategy employed by *A. fumigatus* in the wild.

REFERENCES

1. Dias, D. A., Urban, S. & Roessner, U. A Historical Overview of Natural Products in Drug Discovery. *Metabolites* **2,** 303–336 (2012).

2. Fischbach, M. A. & Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev* **106,** 3468–3496 (2006).

3. Du, L. & Lou, L. PKS and NRPS release mechanisms. *Nat Prod Rep* **27,** 255, (2010).

4. Yadav, G., Gokhale, R. S. & Mohanty, D. Towards Prediction of Metabolic Products of Polyketide Synthases: An In Silico Analysis. *PLoS Comput Biol* **5,** e1000351, (2009).

5. Rubinstein, E. & Keynan, Y. Vancomycin revisited – 60 years later. *Front Public Health.* **2**, 1–7, (2014).

6. Zhang, H., Wang, Y., Wu, J., Skalina, K. & Pfeifer, B. A. Complete Biosynthesis of Erythromycin A and Designed Analogs Using E. colias a Heterologous Host. *Chem Biol* **17**, 1232–1240, (2010).

7. Felnagle, E. A. *et al.* Nonribosomal Peptide Synthetases Involved in the Production of Medically Relevant Natural Products. *Mol Pharm* **5,** 191–211, (2008).

8. Ling, L. L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 1–19, (2015).

9. Baccile, J. A. *et al.* Plant-like biosynthesis of isoquinoline alkaloids in Aspergillus fumigatus. *Nat Chem Biol* **12**, 412-419, (2016).

10. Forseth, R. R. *et al.* Homologous NRPS-like Gene Clusters Mediate Redundant Small-Molecule Biosynthesis in Aspergillus flavus. *Angew Chem Int Ed* **52,** 1590–1594, (2012).

11. Walsh, C. T. The chemical versatility of natural-product assembly lines. *Acc Chem Res* **41**, 4–10, (2008).

12. Bachmann, B. O., Van Lanen, S. G. & Baltz, R. H. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J Ind Microbiol Biotechnol* **41,** 175–184, (2013).

13. Baccile, J. A. *et al.* Modular synthesis of photocleavable peptides using click chemistry. *Tet Let* **53**, 1933–1935, (2012).

14. Zimmermann, M. & Fischbach, M. A. Brief Communication. *Chem Biol* **17**, 925–930, (2010).

15. Iyer, L. M., Abhiman, S., Burroughs, A. M. & Aravind, L. Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol Biosyst* **12**, 1636-1660, (2009).

CHAPTER 6

CONCLUSIONS AND OUTLOOK

Filamentous fungi as a source of new chemistry: While having long been known as prominent producers of small-molecules, filamentous fungi's ability to perform a diverse range of chemical transformations has been largely under appreciated. As evidenced by Chapters 3 and 5, noncanonical NRPS-like enzymes have the potential to produce completely unexpected molecular architectures. In *aspergillus* species alone there remains a significant number of uncharacterized BGCs that include genes encoding for putative NRPS-like enzymes. It is likely that future explorations of these orphan gene clusters will be rewarded with yet more interesting biosynthetic mechanisms and potentially useful compounds. Further, based on genomic sequencing data, NRPS-like enzymes represent only a small portion of the unexplored enzymology in filamentous fungi.

DANS- and LC-MS-based comparative metabolomics: A continued effort to advance DANS- and LC-MS-based comparative metabolomics is critical to the success of the field of chemical biology. To date, no other method has been as efficient at comprehensively annotating metabolic pathways in microorganisms and animals. As evidenced by this dissertation, DANS-based comparative metabolomics is uniquely suited toward the identification of unstable molecules, which represent a particularly important group of metabolites. Moreover, the combination of genetic engineering and DANS provides a

direct feedback loop to understand the biological context of the regulation involved in the anabolism or catabolism of metabolic pathways. This dissertation provides a basis for the reality of completely annotating the microbial metabolome through the systematic analysis of each putative BGC. It will however, be necessary to improve DANS-based comparative metabolomics to eliminate the necessity for extensive human interpretation by creating software packages capable of automatically distinguishing differential features in 2D-NMR spectra.

The metabolome at large: The most profound take away from this dissertation is that despite years of sustained isolation and characterization efforts by natural products chemists the microbial metabolome remains almost completely unknown. It is especially surprising considering that the organisms studied in this dissertation are among the most well-studied microbial species. An extension of this observation is that not only is the microbial metabolome not well understood but all metabolomes are not well understood, especially our own. Comprehensive annotation of the human metabolome has the potential to be as influential as the human genome project and there is no doubt that DANS- and LC-MS-based comparative metabolomics will play a central role.

APPENDIX A

A NONRIBOSOMAL PEPTIDE SYNTHETASE-DERIVED IRON(III) COMPLEX FROM THE PATHOGENIC FUNGUS *ASPERGILLUS FUMIGATUS*

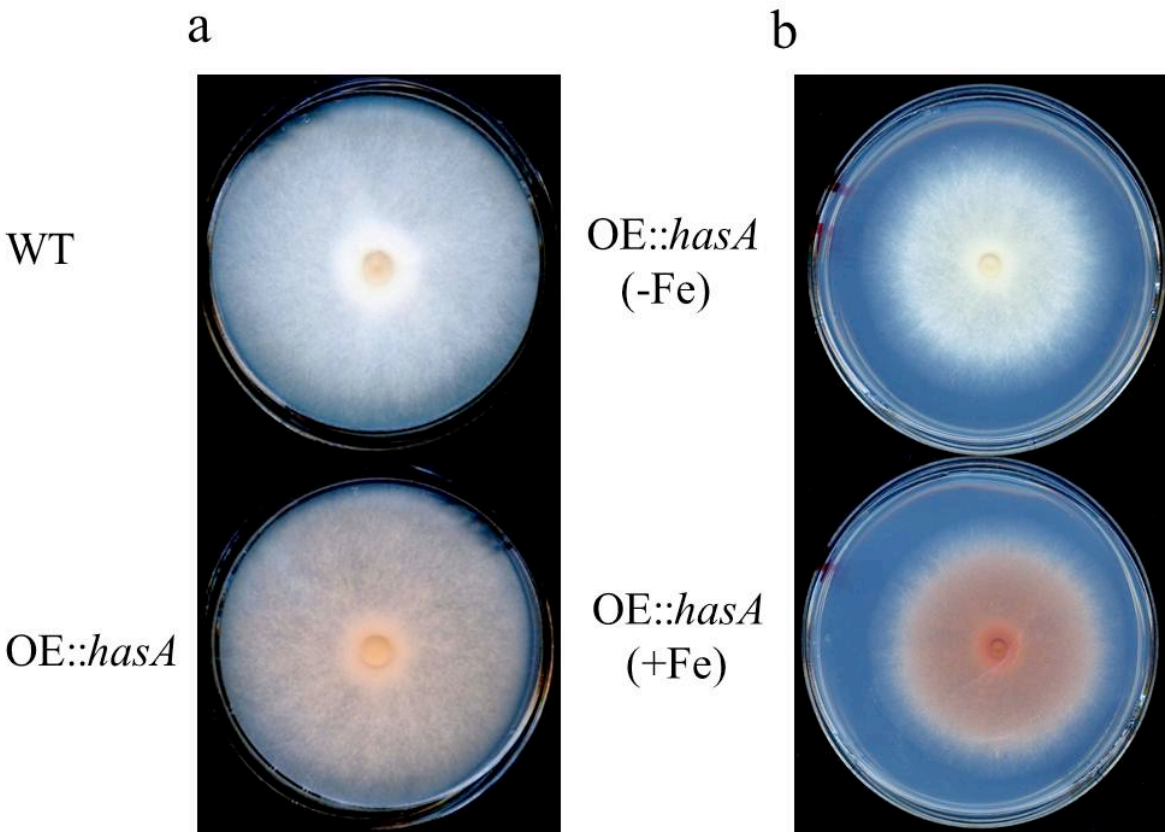


Figure A1. Backside phenotypes of *A. fumigatus* cea10 mutants. a) WT (cea17-1) and OE::*hasA* (TJW109.3) were grown on minimal medium (1.8 μM Fe^{2+}) plates (diameter 55 mm) for 5 days at 37 °C; b) OE::*hasA* (TJW109.3) was grown on minimal medium plates (diameter 90 mm) for 5 days at 37 °C with 1.5 mM Fe^{2+} and without iron.

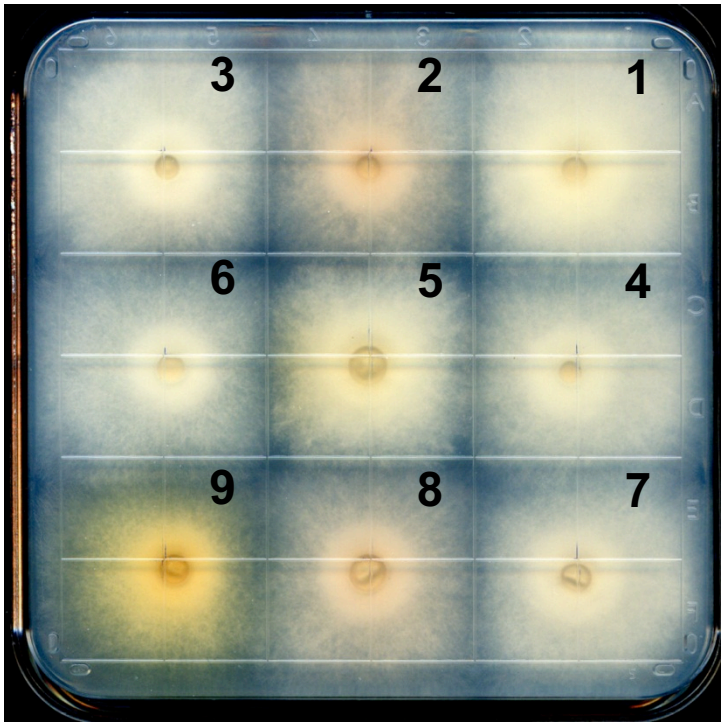


Figure A2. Phenotypes of *has* mutants. Strains were point-inoculated and grown on minimal medium (1.8 μM Fe^{2+}) for 3 days at 37 °C. 1: WT (cea17-1), 2: OE::*hasA* (TJW109.3), 3: OE*::hasA*Δ*hasD* (TWY25.5) , 4: OE*::hasA*Δ*hasE* (TWY26.1), 5: OE*::hasA*Δ*hasC* (TWY29.8), 6: OE*::hasA*Δ*hasH* (TWY27.1), 7: OE*::hasA*Δ*hasB* (TWY30.3), 8: OE*::hasA*Δ*hasF* (TWY31.4), 9: OE*::hasA*Δ*hasG* (TWY28.3).

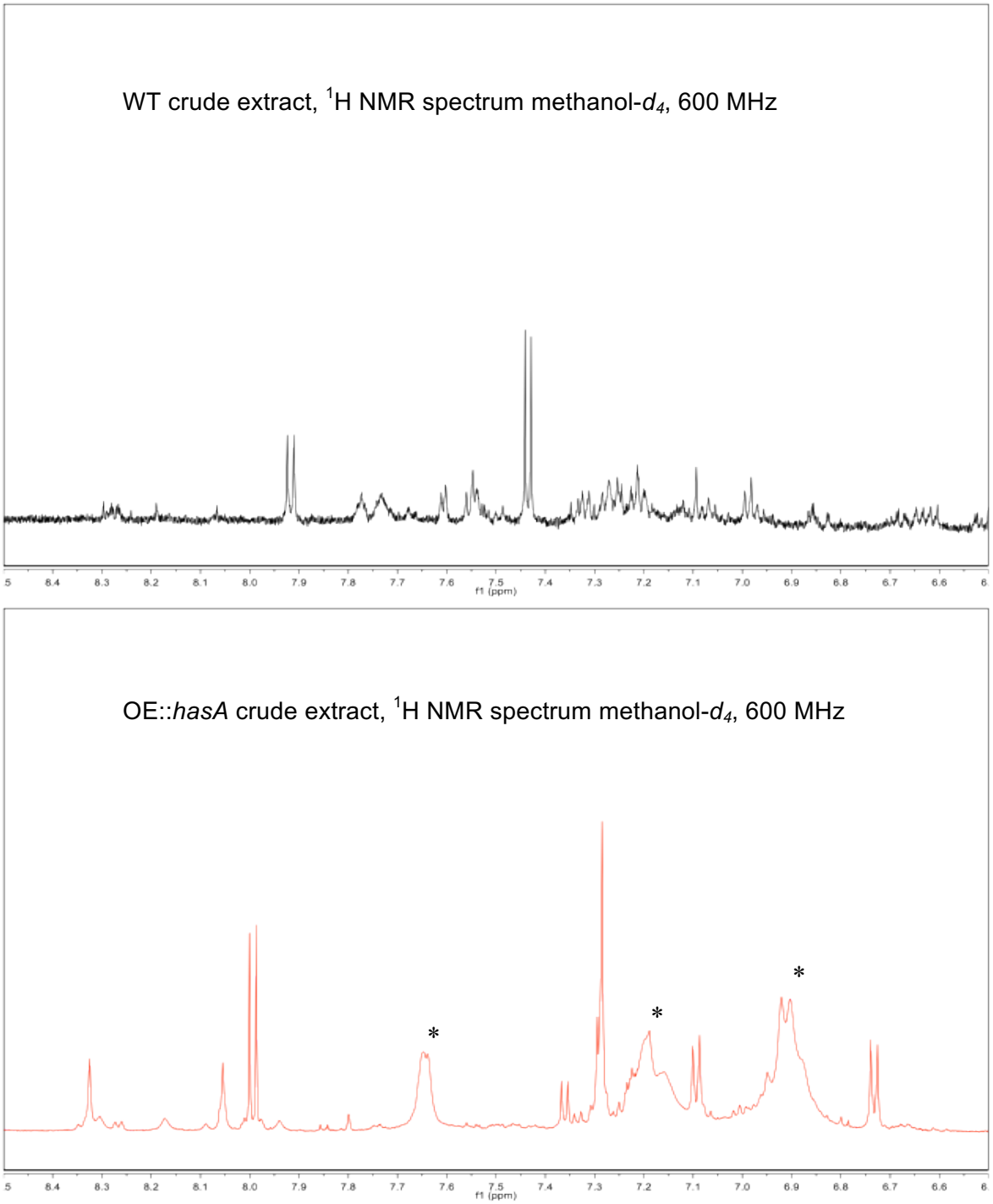


Figure A3. Signals indicative of a *has*-dependent metabolite containing a 7-substituted indole. Comparison of the aromatic region of the ^{1}H NMR spectra from WT metabolite extract (top) and OE::*hasA* extract (bottom). *Indicates signals present in OE::*hasA* and not WT. Broad line shapes due to presence of paramagnetic Fe^{3+} prevents detailed 2D NMR-spectroscopic analysis; however, dqfCOSY spectra revealed weak crosspeaks consistent with a 7-substituted indole. Signals of HAS protons closer to the iron (e.g. the methyl group) are even broader.

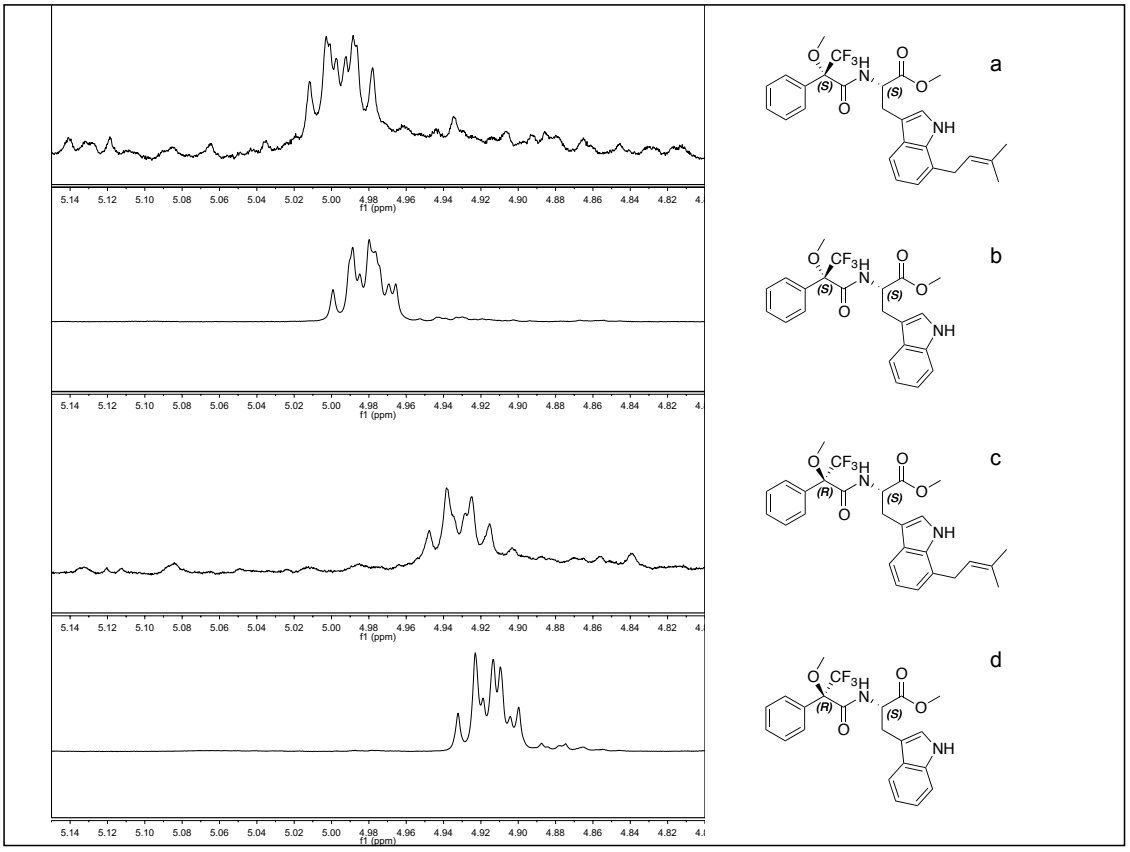


Figure A4. Determination of the stereochemistry of compounds **3** and **4a/b**. (a) Partial ^{1}H NMR spectrum showing the α-proton of the **(*S*)-**MTPA derivative of the methyl ester of **3,** showing an α-proton chemical shift highly similar to that of the **(*S*)-**MTPA derivative of L-trp methyl ester shown in (b). Similarly, the chemical shift of the α-proton of of the **(*R*)-**MTPA derivative of the methyl ester of **3** is highly similar to that of the **(*R*)-**MTPA derivative of L-trp methyl ester shown in (b). The chemical shift values of corresponding MTPA derivatives of the methyl esters of **4a** and **4b** are nearly identical.

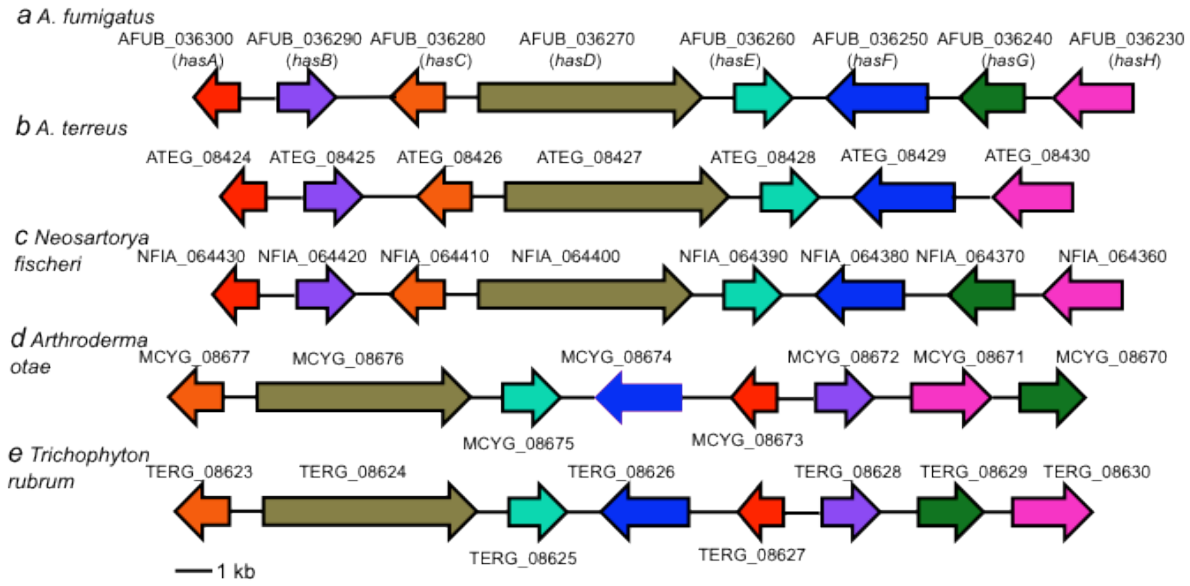


Figure A5. The HAS gene cluster is conserved in a variety of fungal species. a) *A. fumigatus* cea10; b) *A. terreus* NIH2624; c) *Neosartorya fischeri* NRRL181; d) *Arthroderma otae CBS 113480*; e) *Trichophyton rubrum CBS 118892.* Note rearrangement of gene order in the dermatophytic fungi *Arthroderma* and *Trichophyton*.

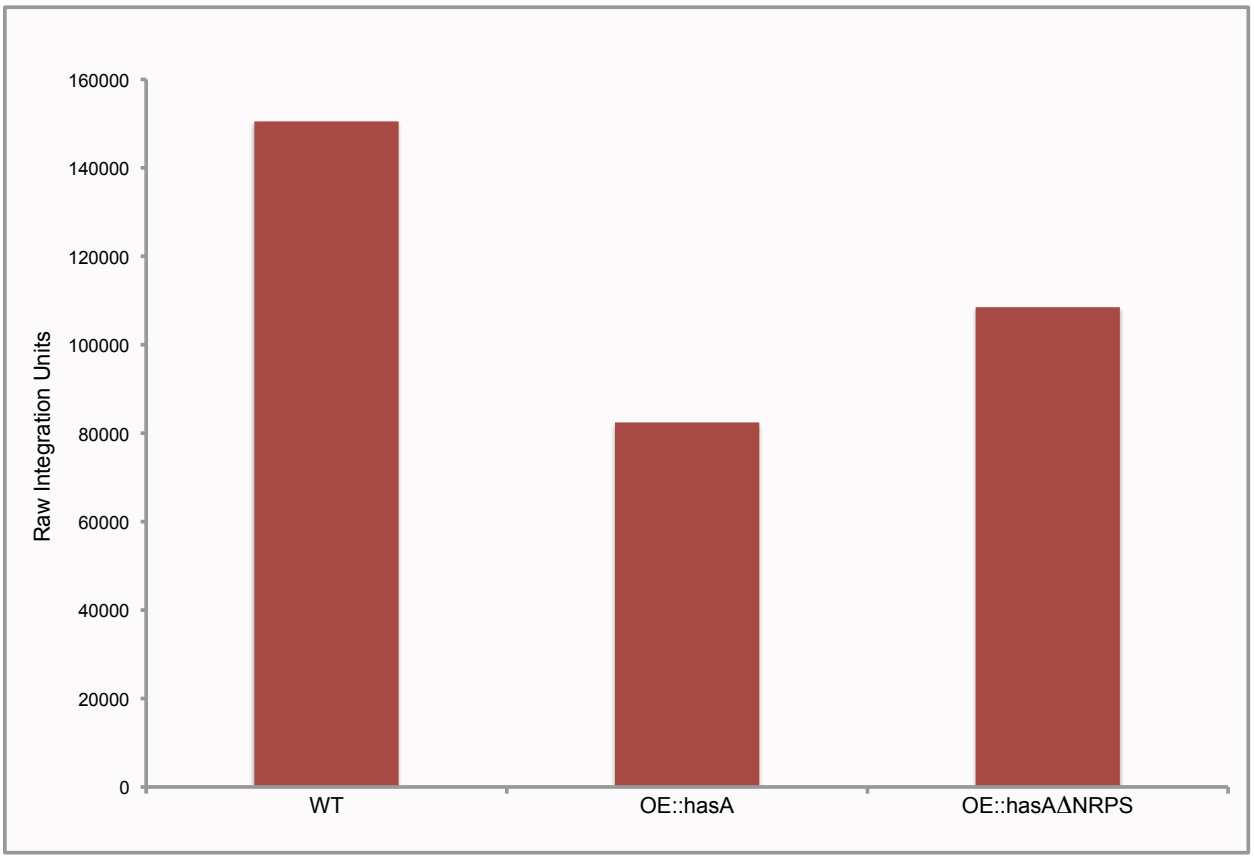


Figure A6. Gliotoxin production in *A. fumigatus*-infected mouse lungs as measured by HPLC-MS. Chromatograms of mouse lung extracts were integrated using MassLynx software. See methods for details. Decreased levels of gliotoxin in OE::*hasA* infected mice lungs indicates HAS and/or HAS associated metabolites may be responsible for the increased virulence observed in this strain.

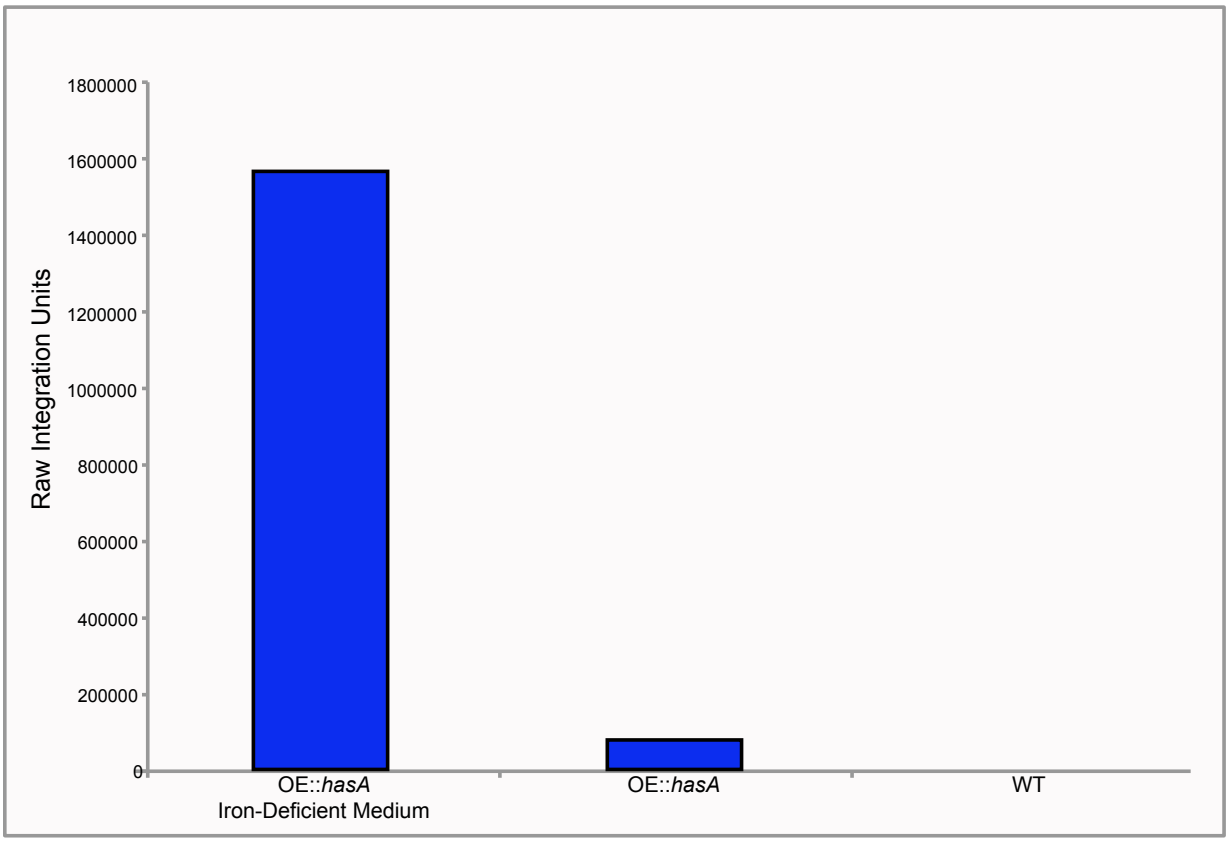


Figure A7. Triacetylfusarinine C (TAFC) production as measured using HPLC-MS. TAFC peaks in the MS chromatograms were integrated using MassLynx software for each sample. Strains were cultured as described in methods. Increased levels of TAFC in OE::*hasA* cultured in iron-deficient medium may indicate increased demand for iron for incorporation into HAS.

Strains, media, and growth conditions. The fungal strains used in this study are listed in Supplementary Table 1. All strains were grown at 37 °C on glucose minimal medium (GMM) [1] and, when appropriate, were supplemented with 0.56 g uracil L^{-1}, 1.26 g uridine L^{-1}, 1.0 g arginine L^{-1} and maintained as glycerol stocks at -80 ºC. *Escherichia coli* strain DH5α was propagated in LB medium with appropriate antibiotics for plasmid DNA.

Gene cloning, plasmid construction, and genetic manipulations. The plasmids utilized in this work are listed in Supplementary Table 1. The oligonucleotide sequences for PCR primers are given in Supplementary Table 2. PCR amplification was carried out on a C1000^{TM} Thermal Cycler from Bio-Rad (Hercules, CA). For creation of the *hasA* overexpression (OE) strain at its native locus, the OE cassette was constructed by using single and double joint PCR procedures.[2] Single joint PCR reactions were set up for fusion of the maker gene *A. parasiticus pyrG* which was amplified from pJMP9.1 and *A. nidulans* glyceraldehyde-3-phosphate dehydrogenase gene *(gpdA*) promoter (1.5 kb), which was amplified from pWY25.16.[3] Then, the 1.03 kb fragment upstream of *hasA* and 1.1 kb fragment containing *hasA* ORF downstream were amplified from genomic DNA of *A. fumigatus* cea17 using designated primers, respectively. These three amplified PCR products were purified with a QIAquick gel extraction kit (Qiagen, Valencia, CA), quantified, and fused using double joint PCR procedures. The final PCR product was amplified with the bottom primer pairs OEgene_5F_for and OEgene3F_rev, confirmed with endonuclease digestion, and purified for fungal transformation. Except for the first round PCR with Pfu Ultra II DNA Polymerases (Agilent, Santa Clara, CA), all PCR steps were performed by using an Expand long template PCR system (Roche, Indianapolis, IN) according to the manufacturer's instructions. Cea17-2 (auxotroph *A. fumigatus* Δ*nkuB, pyrG*-) was used for host strain for transformation.

For the creation of deletion mutants in *OE::hasA* TJW109.3 background, a deletion cassette of each cluster gene (*hasB* to *hasH*) was constructed by using double joint PCR with hygromycin B (*hph*) replacement of target genes. The *hph* gene was amplified by using plasmid pUCH2-8 as template. Five μg of the double-joint PCR

cassette were used to delete each *has* cluster gene by using TJW109.3 (*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1*) as the recipient host. After transformation, transformants were grown on minimal media plates with 100 μg mL^{-1} hygromycin B (GOLDBIO, St. Louis, MO) for screening. To create strains for the pathogenicity assay, we reconstructed OE::*hasA* (TWY22.2) and OE::*hasAΔhasD* (TWY24.121) mutants in *A. fumigatus* 293 (Af293.6) background to avoid any influence of Δ*nkuB* in pathogenicity (all *A. fumigatus* cea10 strains contained Δ*nkuB*). The creation of OE::*hasA* was carried out using ectopic integration of the plasmid pWY47.2. The plasmid pWY47.2 was constructed as follows. Firstly, the open reading frame of *hasA* was amplified by using designated primers WB OETF1_for and WB OETF1_rev. Secondly, plasmid pJMP10.1 (*A. nidulans gpdA(p)* and *A. fumigatus argB* in pBS SK-) was used as template for insertion of *hasA* via quick-change method.[4] The plasmid was then confirmed by sequencing. Ten μg of plasmid pWY47.2 were used for overexpression of *hasA* by using Af293.6 (*A. fumigatus pyrG1, argB1*) as the recipient host. The disruption mutant of *hasD* was created by the replacement of *A. parasiticus* pyrG with TWY22.2 as recipient host. The prototrophic strain of OE::*hasA* was created by transforming plasmid pSA2.4 to TWY22.2. Overexpression and deletion strains were verified by PCR and Southern blot analysis. Three confirmed strains for each mutant were stored at -80 C with 33 % glycerol for future uses.

Nucleic acid analysis. Plasmid preparation, digestion with restriction enzyme, gel electrophoresis, blotting, hybridization, and probe preparation were performed by standard methods.[5] *Aspergillus* DNA for diagnostic PCR was isolated using the previously described method.[6] Sequence data were analyzed using the LASERGENE

software package from DNASTAR.

Northern analysis. 50 mL of liquid GMM were inoculated with 10^6 spores per mL of all appropriate strains in this study and incubated with shaking at 250 rpm at 25 °C. After 44 h, the mycelium was collected and total RNA was extracted by using Isol-RNA Lysis Reagent according to the manufacturer's instructions (5 Prime, Hamburg, Germany).

For the determination of boundaries of the *has* gene cluster, the fragments from *has* gene members used as probes were amplified individually from cea17-1 genomic DNA with appropriate primers (Supplementary Table 2). WT (cea17-1) and OE::*hasA* (TJW109.3) mutants were used for this experiment. For assessment of fumitremogin (*ftm*) cluster expression, WT, OE::*hasA*, OE::*hasA*Δ*hasD* (TWY25.5), and OE::*hasA*Δ*hasE* (TWY26.1) strains were used. About 30 μg of total RNA were used for RNA blot analysis. RNA blots were hybridized with designated DNA fragments, which were generated by PCR using gene-specific primers as shown in Supplementary Table 2. All experiments were performed in duplicate. Detection of signals was carried out with a Storm 860 phosphoimager (Molecular Dynamics, Sunnyvale, CA).

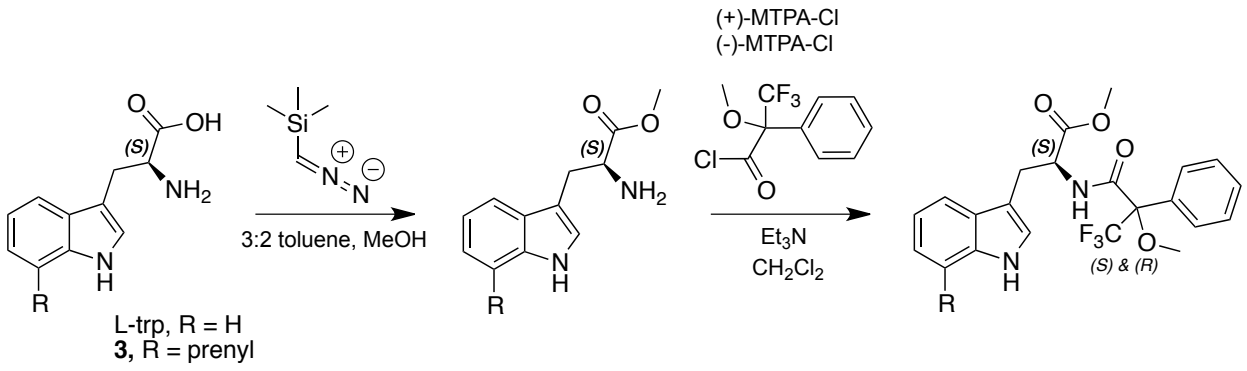
Fermentation and metabolome extraction. *A*. *fumigatus* strains were inoculated (1.0×10^6 spores/mL) into 1 L GMM in a 2 L flask at 25 °C with shaking at 250 rpm. After 3 days, liquid fungal cultures including fungal tissue and media were placed in 1 L Erlenmeyer flasks, frozen using a dry ice acetone bath, and lyophilized. The lyophilized residues were extracted with 500 mL of 10% methanol in ethyl acetate for 3.5 h with vigorously stirring. Extracts were filtered over cotton, evaporated to dryness, and stored in 8 mL vials. (a) Preparation for NMR Analysis: Crude extracts (~30 mg) were

suspended in 0.15 mL of methanol-d_4. The suspension was evaporated to dryness, re-suspended in 0.6 mL of methanol-d_4, centrifuged to remove insoluble materials and the supernatant was subjected to NMR spectroscopic analysis. (b) Preparation for HPLC-MS Analysis: Crude extracts were suspended in 1.5 mL of methanol and centrifuged to remove insoluble materials, and the supernatant was subjected to HPLC-MS analysis.

Analytical methods and equipment overview. (a) NMR Spectroscopy: NMR spectroscopic instrumentation: a Varian INOVA 600 MHz NMR spectrometer (600 MHz ^{1}H reference frequency, 151 MHz for ^{13}C) equipped with an HCN indirect-detection probe. Non-gradient phase-cycled dqfCOSY spectra were acquired using the following parameters: 0.6 s acquisition time, 400-600 complex increments, 16-32 scans per increment. Non-gradient HSQC, HMQC, and HMBC spectra were acquired with these parameters: 0.25 s acquisition time, 200-500 increments, 8-32 scans per increment. ^{1}H, ^{13}C-HMBC spectra were optimized for $J_{H,C}$ = 6 Hz. Susceptibility-matched NMR tubes (Shigemi) were used for sample amounts smaller than 1 mg. NMR spectra were processed and baseline corrected using Varian VNMR and MestreLabs MestReC and MNOVA software packages. (b) Mass spectrometry: High-resolution mass spectrometry was performed on a Waters nanoACQUITY UPLC system equipped with a Waters Acquity UPLC HSS C-18 column (2.1 x 100 mm, 1.8 µm particle diameter) connected to a Xevo G2 QTof Mass Spectrometer operated in electrospray positive (ESI^+) ionization mode. Low-resolution mass spectrometer was performed on an HPLC-MS system equipped with a diode array detector and connected to a Quattro II spectrometer (Micromass/Waters) operated in positive electrospray ionization (ESI^+) mode. Data

acquisition and processing for the HPLC-MS was controlled by Water MassLynx software. (c) Chromatography: Flash chromatography was performed using a Teledyne ISCO CombiFlash system. For semi-preparative HPLC a Supelco Discovery HS C-18 column (25 cm x 10 mm, 5 µm particle diameter) was used. An Agilent Zorbax Eclipse XDB-C18 column (4.6 x 250 mm, 5 µm particle diameter) was used in the HPLC-ESI$^+$-MS A. *fumigatus* mutant profiling analysis.

Chromatographic enrichment and high-resolution MS data for compounds 1-3a/b, 4, 5a/b, 6-8. Compounds were purified via semi-preparative HPLC using a Supelco Discovery HS C-18 column (25 cm x 10 mm, 5 µm particle diameter) a 1:1-mixture of methanol and acetonitrile (organic phase) and 0.1 % acetic acid in water (aqueous phase) as solvents at a flow rate of 3.20 mL/min. A solvent gradient scheme was used, starting at 5% organic for 3 min, followed by a linear increase to 100% organic over 35 min, holding at 100% organic for 15 min, then decreasing back to 5% organic for 1 min and holding at 5% organic for the final 6 min, a total of 1 h. Fractions of interest were evaporated to dryness and then dissolved in 50-300 µL of methanol and analyzed for high-resolution MS using direct infusion onto the Xevo G2 QTof mass spectrometer operated in electrospray positive (ESI$^+$) ionization mode.



Conversion of compound 3 and L-trp into *R*/*S*-MTPA derivatives for NMR spectroscopic analysis. To a solution of an isolated sample of compound **3** (4.0 mg, 0.0147 mmol) in 3:2 toluene/MeOH (0.75 mL) was added (trimethylsilyl)diazomethane (15 µL, 0.0294 mmol) as a 2.0 M diethyl ether solution. After stirring at room temperature for 45 min, the mixture was dried by rotary evaporation. The residue was dissolved in dichloromethane (1.0 mL), split into two vials (0.5 mL each) and treated separately with triethylamine (10.3 µL, 0.0735 mmol) and (+)- or (-)-α-methoxy-α-(trifluoromethyl)phenylacetyl chloride (11 µL, 0.0588 mmol). After stirring for 2 h at room temperature, the mixtures were dried by rotary evaporation and analyzed by ^{1}H NMR spectroscopy (Supplementary Figure 4) without purification. The same procedure was used for the derivatization of L-Trp.

Iron removal from HAS and formation of compound 9. Adapted procedure from Yamamato and co-workers.[7] Isolated HAS or OE::*hasA* extract (110 mg) was suspended in acetone and any insoluble materials were filtered off. To the filtrate was added 0.2 M NaOH (0.75 mL) dropwise with stirring, the resulting brown precipitate was filtered off, and the acetone was removed from the filtrate by rotary evaporation. The aqueous residue was neutralized by the addition of 0.2 M HCl (0.75 mL), and the resulting precipitate (containing **9**) was filtered off, washed with water (2x) and then dried by rotary evaporation. For spectroscopic data of **9**, see Table S9.

Animal model of *A. fumigatus* infection. Virulence of *A. fumigatus* wild type (Af293), TWY32.1 and TWY24.121 strains was assessed in the neutropenic murine pulmonary

model. Briefly, female Swiss ICR mice (Harlan Sprague Dawley) weighing about 18-20 g were immunosuppressed via administration of cyclophosphamide by separate subcutaneous injections, one at 4 days (200 mg kg^{-1} of body weight) and the other at 1 day (200 mg kg^{-1}) before infection. The second treatment includes administration of cortisone acetate at a dose of 250 mg kg^{-1} by separate subcutaneous injection. Anesthetized mice (10 mice/fungal strain) were infected by nasal instillation of 50 µL of 1 x 10^7 conidia mL^{-1} (day 1) and monitored three times daily for 7 days post-infection. All surviving mice were sacrificed at day 7. The Keller animal protocol number is M02468.

Murine lung tissue extraction and HPLC-MS analysis. Lung tissue samples from infected mice with all three fungal strains were extracted and subsequently analyzed by HPLC-MS as follows. (a) Extraction: mice lungs were lyophilized, placed in 125 mL Erlenmeyer flasks, and extracted with 75 mL of 10% methanol in ethyl acetate in the presence of ~2 g of pre-washed chromatography grade sand for 10 h. The resulting pulverized lung material was filtered over pre-washed cotton and dried by rotary evaporation. (b) HPLC-MS Analysis: the crude mice lung extracts were analyzed by HPLC-MS using a 1:1 mixture of methanol:acetonitrile (organic phase) and 0.1 % acetic acid in water (aqueous phase) as solvents at a flow rate of 1.0 mL/min. A solvent gradient scheme was used starting at 5% organic for 3 min, followed by a linear increase to 100% organic over 35 min, holding at 100% organic for 15 min, then decreasing back to 5% organic for 1 min and holding at 5% organic for the final 6 min. The Quattro II mass spectrometer was operated in single ion recording (SIR) mode for the masses listed in Table S4, corresponding to compounds **1,3-5a/b,6,7,9**. Retention times for these compounds were established using purified standards.

Detection of gliotoxin in mice lung samples. Lung tissue extracts described in the previous section were analyzed by HPLC-MS using the same chromatographic method as before but the Quattro II mass spectrometer was operated in single ion recording (SIR) mode for mass 263.1 (gliotoxin $[M\text{-}S_2+H]^+$), which was the most abundant ion for gliotoxin in ESI^+ mode.

Table A1. Fungal strains and plasmids used in this study

Strain/plasmid	Description	Reference
***Aspergillus fumigatus* 293 background strains**		
Af293	Wild type	[8]
Af293.1	*A. fumigatus pyrG1*	[8]
Af293.6	*A. fumigatus pyrG1, argB1*	[8]
TWY22.2	*gpdA(p)::hasA::argB1::argB, pyrG1*	This study
TWY24.121	*gpdA(p)::hasA::argB1::argB, ΔhasD::AppyrG, pyrG1*	This study
TWY32.1	*gpdA(p)::hasA::argB1::argB, pyrG1::AppyrG*	This study
***Aspergillus fumigatus* cea10 background strains**		
Cea17-1	Prototroph *A.fumigatus* ΔnkuB, pyrG+	[9]
Cea17-2	Auxotroph *A.fumigatus* ΔnkuB, pyrG-	[9]
TJW109.3	*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1*	This study
TWY25.5	*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1, ΔhasD::hph*	This study
TWY26.1	*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1, ΔhasE::hph*	This study
TWY27.1	*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1, ΔhasH::hph*	This study
TWY28.3	*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1, ΔhasG::hph*	This study
TWY29.8	*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1, ΔhasC::hph*	This study
TWY30.3	*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1, ΔhasB::hph*	This study
TWY31.4	*A.ppyrG::gpdA(p):: hasA, ΔnkuB, pyrG1, ΔhasF::hph*	This study
Plasmids		
pJMP9.1	*gpdA(p)::A. para pyrG* cassette	[10]

pJMP10.1	*A. nidulans* gpdA(p) and *A. fumigatus* argB in pBS SK-	Palmer, J
pSA2.4	*A.fumigatus pyrG* 1.97kb in TOPOpCR2.1	[11]
pWY47.2	gpdA:: hasA in pJMP10.1	This study
pUCH2-8	Hygromycin resistance cassette	[12]
pWY25.16	*A.fumigatus pyroA::gpdA* in pGEMT easy vector	[3]

pXX = plasmid, TXX = original transformant

Table A2. PCR primer sets used in this study

Name of the primer	Oligonucleotide sequence (5'-3')	Uses
363005'F	AAGGAAGATCACCGTCAACGCG	Amplification of OEhasA 5' flanking region (cea10)
363005'R	ccaattcgccctatagtgagtcgtattacgAGTGGTGC CAAGAGTTCTCAACCAAGGCG	
363003'F	gctaccccgcttgagcagacatcaccatgACTTCTAC ACTTCCTTATTTGACCAGTCC	Amplification of OEhasA 3' flanking region (cea10)
363003'R	TTCCTTGGTGACTTTGTAGTCCC	
36300conf	GTCTACGGATCCATTCACCAGG	OEhasA screening
WB OETF1_for WB OETF1_rev	cagctaccccgcttgagcagacatcaccATGACTTC TACACTTCCTTATTTG gatggttctctcctgcacatagccctcggCTGGTGGG TAAAAGACCAAG	*hasA* amplification for OE::hasA cassette (Af293)
TFseq_for	GCTACATCCATACTCCATCC	sequencing and screening
TFseq_rev	GCGCGCAATTAACCCTCAC	sequencing
TF screening_for	CCTTATTCGTTGACCTAGCTG	OE::hasA screening
TF1 screening_rev	CGATGTTCAGGTTTCCCAGC	
NRPS Nor_for	ATGAACCTGACTCTCACTAC	hasD northern probe
NRPS Nor_rev	CGTCAAAAGCGATGGAACAG	
DMAT Nor_for	ATGTCCATCGGAGCCGAG	hasE northern probe
DMAT Nor_rev	GCGCTAGATGCAGAATGG	
TF2 Nor_for	GTTTGTCGGATCTCAGGG	hasF northern probe
TF2 Nor_rev	CTTGCCAATCCCGGGAAG	
P450 Nor_for	CTCATTTGGCGTCAGTTCAC	hasH northern probe
P450 Nor_rev	GATAACGGACCCCAAAGACG	
MT Nor_for	CTACATCAACTCAGACTCCC	hasC northern probe
MT Nor_rev	CAGAGCTTTTCTACAACACC	

TF1 Nor_for	GAAGATCAGCTCGATTGCGG	hasA northern probe
TF1 Nor_rev	GTTTCCGTTGCACGTCCAAG	
36290 Nor_for	CGCAGGTGGAGTGATTTATCC	hasB northern probe
36290 Nor_rev	GAGAATCATCGTTGCTCCAGC	
FAD Nor_for	CAATTCCTTGCACCACACC	hasG northern probe
FAD Nor_rev	CCCACTGGCTGACTTTGAG	
36310 Nor_for	GGTCACAGAAAATGGAGAG	36310 northern probe
36310 Nor_rev	CTTCGAGGCTAATCTCATG	
36320 Nor_for	CATGGGTAGCCAGAAGCG	36320 northern probe
36320 Nor_rev	CATACTTCCTCGTCCACTC	
36220 Nor_for	GTCTTCCTGGGATATGTC	36220 northern probe
36220 Nor_rev	CTTCCGAAACTGACGCTG	
KONRPS5F_for	GCCATTACCATTAGCACGTG	Amplification of KOhasD 5' flanking region (hph)
KONRPS5F_rev	GCGCATTGTTAGATTTCATACACGGTGC CTGGCCTTCCAACAAGTATGCTC	
KONRPS3F_for	GTCAGGCCATTTTCATATGGCAATGCGC AGCAGTCATGGTATGTTGACCG	Amplification of KOhasD 3' flanking region (hph)
KONRPS3F_rev	CCTTTGTGTTCTCGTCTGTC	
KONRPS5F_rev (pyrG)	ccaattcgccctatagtgagtcgtattacgGCCTTCCA ACAAGTATGCTC	Amplification of KOhasD 5' flanking region (AppyrG)
KONRPS3F_for (pyrG)	cagcttatcgatgataagctgtcaaacatgagCAGTCA TGGTATGTTGACCG	Amplification of KOhasD 3' flanking region (AppyrG)
KODMATS5F_for	GAAGATATCGCACTGAGGCTG	Amplification of KOhasE 5' flanking region
KODMATS5F_rev	GCGCATTGTTAGATTTCATACACGGTGC CTGGTTCACTGCTGTCCTGAGTC	
KODMATS3F_for	GTCAGGCCATTTTCATATGGCAATGCGC AGGCACTGATGGTGTTCATTTGAG	Amplification of KOhasE 3' flanking

KODMATS3F_rev	CAATACTTGCAATACTGCGACC	region
KOMT5F_for	CATTGCCAACCTTCTCTGCTC	Amplification of KOhasC 5' flanking region
KOMT5F_rev	GCGCATTGTTAGATTTCATACACGGTGC CTGGTTGGGAAGTGTCGTTATGC	
KOMT3F_for	GTCAGGCCATTTTCATATGGCAATGCGC AGCATTCTGAATCTCTGCCGG	Amplification of KOhasC 3' flanking region
KOMT3F_rev	GGATACTTGGAGTCAATGGC	
KOP450 5F_for	GAAAGACCAGACAGCTGTGC	Amplification of KOhasH 5' flanking region
KOP4505F_rev	GCGCATTGTTAGATTTCATACACGGTGC CTGCTGCACTAGATGGACGCTTG	
KOP4503F_for	GTCAGGCCATTTTCATATGGCAATGCGC AGGCATATTTGTAGTCGTGGTCC	Amplification of KOhasH 3' flanking region
KOP450 3F_rev	GGGATTTTGCATCCCAGGTTG	
KO362900 5F_for	GTGACGATGCTTAGCATCTTC	Amplification of KO36290 5' flanking region
KO36290 5F_rev	GCGCATTGTTAGATTTCATACACGGTGC CTGGAGAAGCAATTACCGTAGAGG	
KO36290 3F_for	GTCAGGCCATTTTCATATGGCAATGCGC AGCGATTCTCTTGCAACCTTGTG	Amplification of KO36290 3' flanking region
KO36290 3F_rev	CTTCCACAAAGTCACCGGTAC	
KOTF2 5F_for	CAACGACATGTTTGTTGGCC	Amplification of KOhasF 5' flanking region
KOTF25F_rev	GCGCATTGTTAGATTTCATACACGGTGC CTGCTGCACATCTACCTCATGCTC	
KOTF2 3F_for	GTCAGGCCATTTTCATATGGCAATGCGC AGGTGACAGCCTTGGAATGATG	Amplification of KOhasF 3' flanking region
KOTF2 3F_rev	GCTCCCGTCTATGACAGATC	
KOFAD5F_for	CAACGGCAGAAAGATTGTACC	Amplification of KOhasG 5' flanking region
KOFAD5F_rev	GCGCATTGTTAGATTTCATACACGGTGC CTGGGATGGATCTGAGTTGGTTG	

KOFAD3F_for	GTCAGGCCATTTTCATATGGCAATGCGC AGGGTGTTCTAGCTTCGGCTTAC	Amplification of KOhasG 3' flanking region
KOFAD3F_rev	CTAGGTGACTGACACGATTGC	
WB hyg_for	CAGGCACCGTGTATGAAATCTAAC	Hygromycin amplification
WB hyg_rev	CTGCGCATTGCCATATGAAAATGGC	
WB hygint_rev	GTCCGAGGGCAAAGGAATAG	screening primer for KO mutant
KONRPS_Scr_for	CCACTGGTATCTACCCTACATC	KOhasD screening
KODMATS_Scr_for	GTTCACTGAGATGGCGACAC	KOhasE screening
KOP450_Scr_for	CTCAAAGTCAGCCAGTGGG	KOhasH screening
KO12900_Scr_for	GATCGGAATTGCTCGGATGG	KOhasB screening
KOTF2_Scr_for	GCAGTTCATGCACTCATACG	KOhasF screening
KOFAD_Scr_for	CTAGATTGATCTCCCAGCAGC	KOhasG screening
Nor_ftmA_for	GTTGACGGAACATTCGGCTG	ftmA northern probe
Nor_ftmA_rev	GCATTCACGCTTCTGGATCC	
Nor_ftmB_for	GAAACGGGGATCATTGCG	ftmB northern probe
Nor_ftmB_rev	GAAGAACGAACACGCGTCG	
Nor_ftmC_for	CATGGTGTCTGTCGCAAAC	ftmC northern probe
Nor_ftmC_rev	GGTTATCCTTCCGTCTGAG	
Nor_ftmD_for	GACACAGGCAGTCGACATCG	ftmD northern probe
Nor_ftmD_rev	CGACCAACGATTGAACCGTGC	
Nor_ftmE_for	CATCATCGTCCTCCCCATTC	ftmE northern probe
Nor_ftmE_rev	GCCACAGTCAGAAAGACACTG	
Nor_ftmF_for	GACTCCAAGCCACAACTCC	ftmF northern probe
Nor_ftmF_rev	CTGTCCACCGATGACTGCC	
Nor_ftmG_for	GATCACAGCGGGTTCAAG	ftmG northern probe
Nor_ftmG_rev	CGTACTGGCCTTGATTGC	

Nor_ftmH_for	CCAACTGAGATTTCGTGC	ftmH northern probe
Nor_ftmH_rev	GAAGAAGTAGGCTTTGGCG	
Nor_ftmO_for	GTACTCATCTAGTCTCCGCG	ftmO northern probe
Nor_ftmO_rev	GAAGCTCCTGTACGATAGC	
actin_for	GTATGTCGGTGATGAGGCAC	actin northern probe
actin_rev	CCGTAGAGATCCTTACGGAC	

Table A3. Genes and enzymes predicted through deep bioinformatic annotation of the putative *has*-cluster in fungi

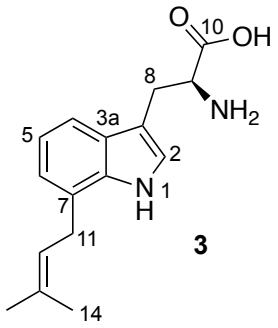
Protein	Identity* (%) to genes of					putative function
A. fumigatus cea10	*A. fumigatus* 293	*N. fischeri*	*A. terreus*	*Arthroderma otae*	*Trichophyton rubrum*	
HasA (AFUB_036300)	AFUA_3G12890	NFIA_064430 (94)	ATEG_08424 (74)	MCYG_08673 (55)	TERG_08627 (52)	C6 transcription factor
HasB (AFUB_036290)	AFUA_3G12900	NFIA_064420 (96)	ATEG_08425 (87)	MCYG_08672 (85)	TERG_08628 (84)	MFS transporter
HasC (AFUB_036280)	AFUA_3G12910	NFIA_064410 (97)	ATEG_08426 (88)	MCYG_08677 (81)	TERG_08623 (82)	methyltransferase
HasD (AFUB_036270)	AFUA_3G12920	NFIA_064400 (94)	ATEG_08427 (76)	MCYG_08676 (62)	TERG_08624 (56)	non-ribosomal peptide synthetase
HasE (AFUB_036260)	AFUA_3G12930	NFIA_064390 (95)	ATEG_08428 (83)	MCYG_08675 (71)	TERG_08625 (70)	7-DMATS [13]
HasF (AFUB_036250)	AFUA_3G12940	NFIA_064380 (96)	ATEG_08429 (83)	MCYG_08674 (71)	TERG_08626 (67)	C6 transcription factor
HasG (AFUB_036240)	AFUA_3G12950	NFIA_064370 (96)	-	MCYG_08670 (70)	TERG_08630 (65)	FAD binding domain protein
HasH (AFUB_036230)	AFUA_3G12960	NFIA_064360 (92)	ATEG_08430 (85)	MCYG_08671 (67)	TERG_08629 (66)	cytochrome P450 monooxygenase

*Percentage similarity values of the *A. fumigatus* cea10 ORFs to those of *N. fischeri*, *A. terreus*, *Arthroderma otae* and *Trichophyton rubrum* clusters, are given in parentheses.

Table A4. HR-HPLC-MS data of reported compounds

Compound	HR-ESI+ Observed (m/z)	Ion	Calculated Ion Formula	Calculated m/z	Retention time [min]
1	1129.3751	$[M + Na]^+$	$C_{60}H_{60}FeN_9NaO_9$	1129.3761	46.87
2	1135.4259	$[M + Na]^+$	$C_{60}H_{66}FeN_9NaO_9$	1135.4231	48.82
3	273.1605	$[M + H]^+$	$C_{16}H_{21}N_2O_2$	273.1603	24.17
4a	271.1451	$[M + H]^+$	$C_{16}H_{19}N_2O_2$	271.1447	24.01
4b	271.1452	$[M + H]^+$	$C_{16}H_{19}N_2O_2$	271.1447	23.15
5a	372.1691	$[M + Na]^+$	$C_{21}H_{23}N_3NaO_2$	372.1688	29.20
5b	402.1787	$[M + Na]^+$	$C_{22}H_{25}N_3NaO_3$	402.1294	28.56
6	306.1224	$[M + Na]^+$	$C_{16}H_{17}N_3NaO_2$	306.1218	19.79
7	348.1694	$[M + Na]^+$	$C_{19}H_{23}N_3NaO_2$	348.1688	26.90
8	364.1620	$[M + Na]^+$	$C_{19}H_{23}N_3NaO_3$	364.1637	27.56
9	352.1662	$[M + H]^+$	$C_{20}H_{22}N_3O_3$	352.1661	-

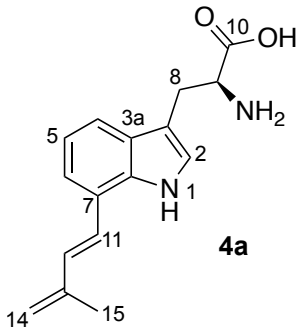
Table A5. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for compound **3** in methanol-d_4. Chemical shifts were referenced to δ($C\underline{H}D_2OD$) = 3.31 ppm and δ($^{13}\underline{C}HD_2OD$) = 49.0 ppm.^{13}C chemical shifts were determined via HMBC and HMQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



3

Position	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1		1-NH		
2	125.05	2-H	7.20	3, 3a, 7a
3	109.87			
3a	128.38			
4	117.09	4-H	7.54 ($J_{5\text{-H}}$ = 8)	6, 7a
5	120.29	5-H	6.99 ($J_{4\text{-H, }6\text{-H}}$ = 8)	3a, 7
6	121.79	6-H	6.92 ($J_{5\text{-H}}$ = 8)	4, 7a, 11
7	125.79			
7a	137.03			
8	28.21	8-H_a	3.52 ($J_{8\text{-Hb}}$ = 15, $J_{9\text{-H}}$ = 5)	2, 3, 3a, 9
		8-H_b	3.15 ($J_{8\text{-Ha}}$ = 15, $J_{9\text{-H}}$ = 9)	2, 3, 9
9	56.54	9-H	3.85 ($J_{8\text{-Ha}}$ = 5, $J_{8\text{-Hb}}$ = 9)	3, 8, 10
10	174.30			
11	30.21	11-CH_2	3.55 ($J_{12\text{-H}}$ = 7)	6, 7, 7a, 12, 13
12	123.21	12-H	5.43 ($J_{11\text{-H}}$ = 7, $J_{14a/b\text{-CH3}}$ < 1)	
13	133.67			
14a	25.88	14a-CH_3	1.76	12, 13, 14b
14b	17.90	14b-CH_3	1.76	12, 13, 14a

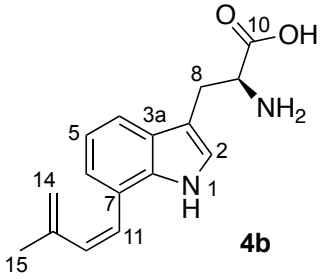
Table A6. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for compound **4a** in methanol-d_4. Chemical shifts were referenced to δ($C\underline{H}D_2OD$) = 3.31 ppm and δ($^{13}\underline{C}HD_2OD$) = 49.0 ppm.^{13}C chemical shifts were determined via HMBC and HMQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



4a

Position	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1		1-NH		
2	125.13	2-H	7.24	3, 3a, 7a
3	109.59			
3a	129.12			
4	118.43	4-H	7.62 ($J_{5\text{-}H}$ = 8)	3, 6, 7a
5	120.34	5-H	7.07 ($J_{4\text{-}H,\ 6\text{-}H}$ = 8)	3a, 7
6	119.40	6-H	7.35 ($J_{5\text{-}H}$ = 8)	4, 7a, 11
7	122.40			
7a	136.23			
8	28.21	8-H_a	3.52 ($J_{8\text{-}Hb}$ = 15, $J_{9\text{-}H}$ = 5)	2, 3, 3a, 9,10
		8-H_b	3.15 ($J_{8\text{-}Ha}$ = 15, $J_{9\text{-}H}$ = 9)	2, 3, 3a, 9, 10
9	56.54	9-H	3.85 ($J_{8\text{-}Ha}$ = 5, $J_{8\text{-}Hb}$ = 9)	3, 8, 10
10	174.30			
11	124.93	11-H	6.99 ($J_{12\text{-}H}$ = 15)	13
12	132.87	12-H	7.04 ($J_{11\text{-}H}$ = 15)	7, 13
13	143.60			
14	117.12	14-H	5.16 ($J_{14\text{-}H'}$ = 4)	12, 15
		14-H'	5.08 ($J_{14\text{-}H}$ = 4)	12, 15
15	18.83	15-CH_3	2.06	12, 13, 14

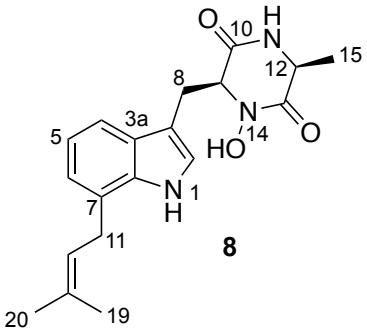
Table A7. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for compound **4b** in methanol-d_4. Chemical shifts were referenced to δ($C\underline{H}D_2OD$) = 3.31 ppm and δ($^{13}\underline{C}HD_2OD$) = 49.0 ppm.^{13}C chemical shifts were determined via HMBC and HMQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



4b

Position	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1		1-NH		
2	125.16	2-H	7.20	3, 3a, 7a
3	109.59			
3a	128.07			
4	118.53	4-H	7.62 ($J_{5\text{-H}}$ = 8)	6, 7a
5	119.70	5-H	7.01 ($J_{4\text{-H, 6-H}}$ = 8)	3a, 7
6	123.19	6-H	7.05 ($J_{5\text{-H}}$ = 8)	4, 7a, 11
7	122.87			
7a	135.97			
8	28.40	8-H_a	3.52 ($J_{8\text{-Hb}}$ = 15, $J_{9\text{-H}}$ = 5)	2, 3, 3a, 9,10
		8-H_b	3.15 ($J_{8\text{-Ha}}$ = 15, $J_{9\text{-H}}$ = 9)	2, 3, 3a, 9, 10
9	56.54	9-H	3.85 ($J_{8\text{-Ha}}$ = 5, $J_{8\text{-Hb}}$ = 9)	3, 8, 10
10	174.30			
11	126.04	11-H	6.65 ($J_{12\text{-H}}$ = 12)	7, 13
12	134.73	12-H	6.40 ($J_{11\text{-H}}$ = 12)	7, 13, 14, 15
13	143.60			
14	118.12	14-H	4.95 ($J_{14\text{-H'}}$ = 4)	12, 15
		14-H'	5.02 ($J_{14\text{-H}}$ = 4)	12, 15
15	21.65	15-CH_3	1.56	12, 13, 14

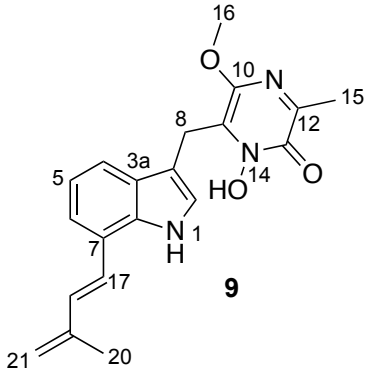
Table A8. ^{1}H (600 MHz) and NMR spectroscopic data for compound **8** in chloroform-*d.* Chemical shifts were referenced to δ($C\underline{H}Cl_3$) = 7.26 ppm ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. Carbon chemical shifts matched published data;[14] however, our proton chemical shifts differed slightly from published values and therefore are listed below.



8

Position	Proton	δH(J_{HH}[Hz])
1	1-NH	8.08
2	2-H	7.03
3		
3a		
4	4-H	7.52 ($J_{5\text{-H}}$ = 8)
5	5-H	7.05 ($J_{4\text{-H, }6\text{-H}}$ = 8)
6	6-H	6.97 ($J_{5\text{-H}}$ = 7)
7		
7a		
8	8-H_a	3.58 ($J_{8\text{-Hb}}$ = 15, $J_{9\text{-H}}$ = 4)
	8-H_b	3.59 ($J_{8\text{-Ha}}$ =15, $J_{9\text{-H}}$ = 3)
9	9-H	4.69 ($J_{8\text{-Ha}}$ = 4, $J_{8\text{-Hb}}$ = 3)
10		
11	11-NH	5.46
12	12-H	3.90 ($J_{15\text{-CH3}}$ = 7)
13		
14		
15	15-CH_3	0.26 ($J_{12\text{-H}}$ = 7)
16	16-CH_2	3.52 ($J_{17\text{-H}}$ = 7)
17	17-H	5.29
18		
19	19-CH_3	1.74
20	20-CH_3	1.79

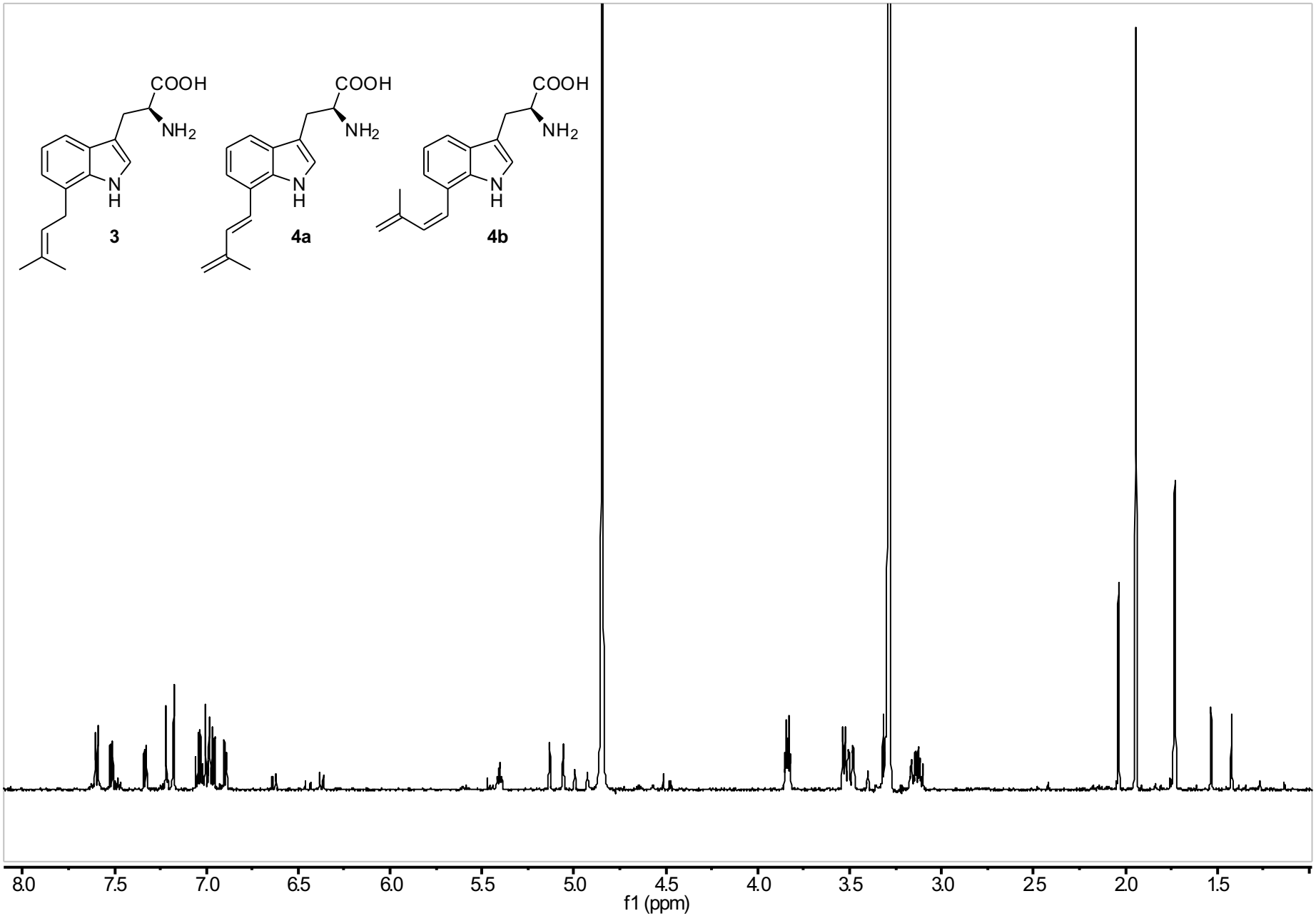
Table A9. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for compound **9** in chloroform-*d*. Chemical shifts were referenced to δ($C\underline{H}Cl_3$) = 7.26 ppm and δ($^{13}\underline{C}DCl_3$) = 77.16 ppm.^{13}C chemical shifts were determined via HMBC and HMQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



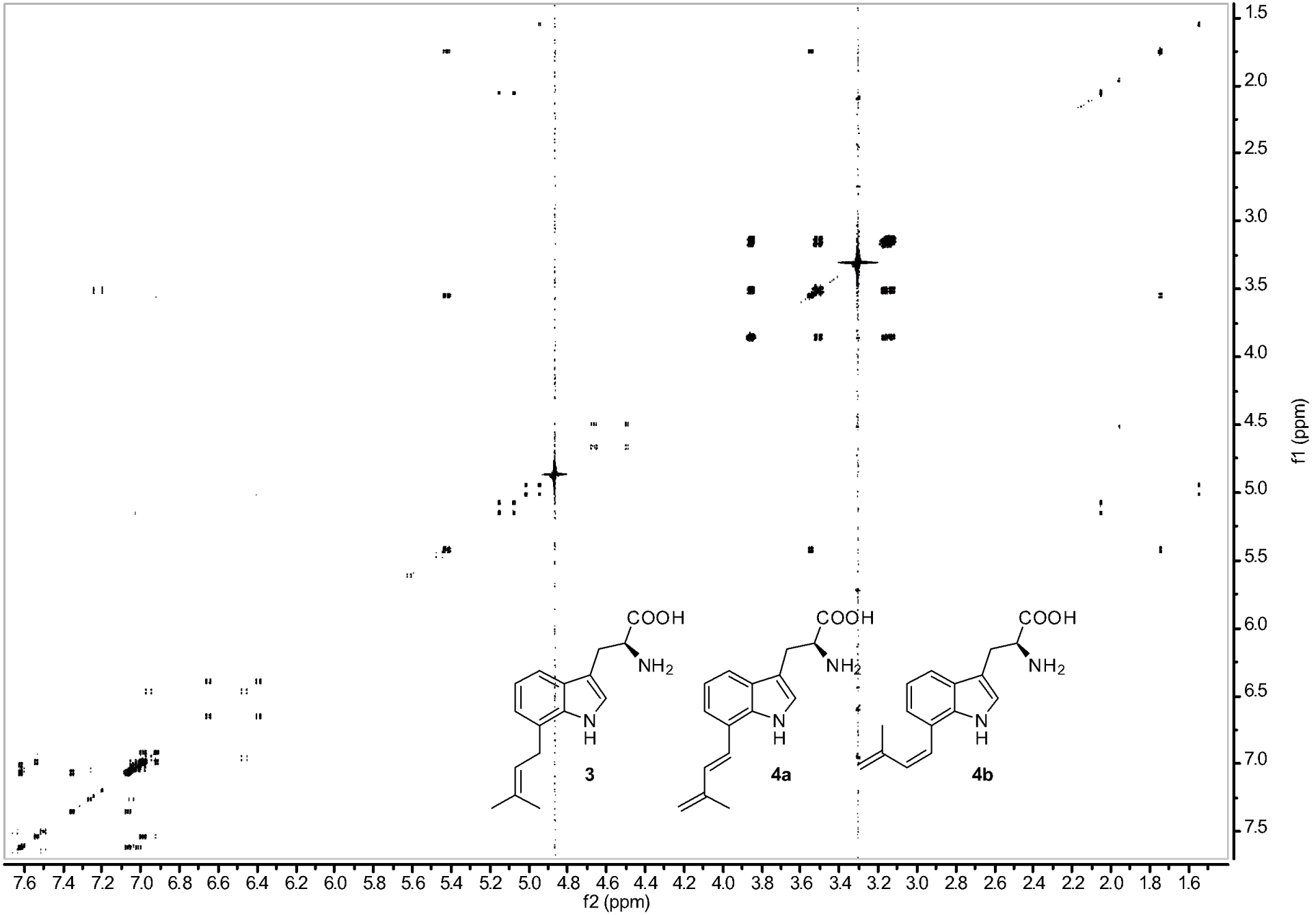
9

Position	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1		1-NH	8.30	
2	123.82	2-H	7.31	3, 3a, 7a
3	110.61			
3a	127.76			
4	118.72	4-H	7.75 ($J_{5\text{-H}}$ = 8)	3, 6, 7a
5	119.84	5-H	7.12 ($J_{4\text{-H, }6\text{-H}}$ = 8)	3a, 7
6	120.33	6-H	7.26 ($J_{5\text{-H}}$ = 8)	4, 7a, 17
7	121.32			
7a	133.58			
8	20.00	8-CH_2	4.36	2, 3, 3a, 9, 10
9	127.39			
10	149.80			
11				
12	134.39			
13	146.74			
14				
15	18.38	15-CH_3	2.42	12, 13
16	54.17	16-OCH_3	3.99	10
17	124.57	17-H	6.73 ($J_{18\text{-H}}$ = 16)	6, 7a, 19
18	132.64	18-H	6.91 ($J_{17\text{-H}}$ = 16)	7, 19, 20, 21
19	141.82			
20	18.53	20-CH_3	2.01	19, 21
21	117.35	21-H	5.09	18, 19, 20
21		21-H’	5.12	18, 19, 20

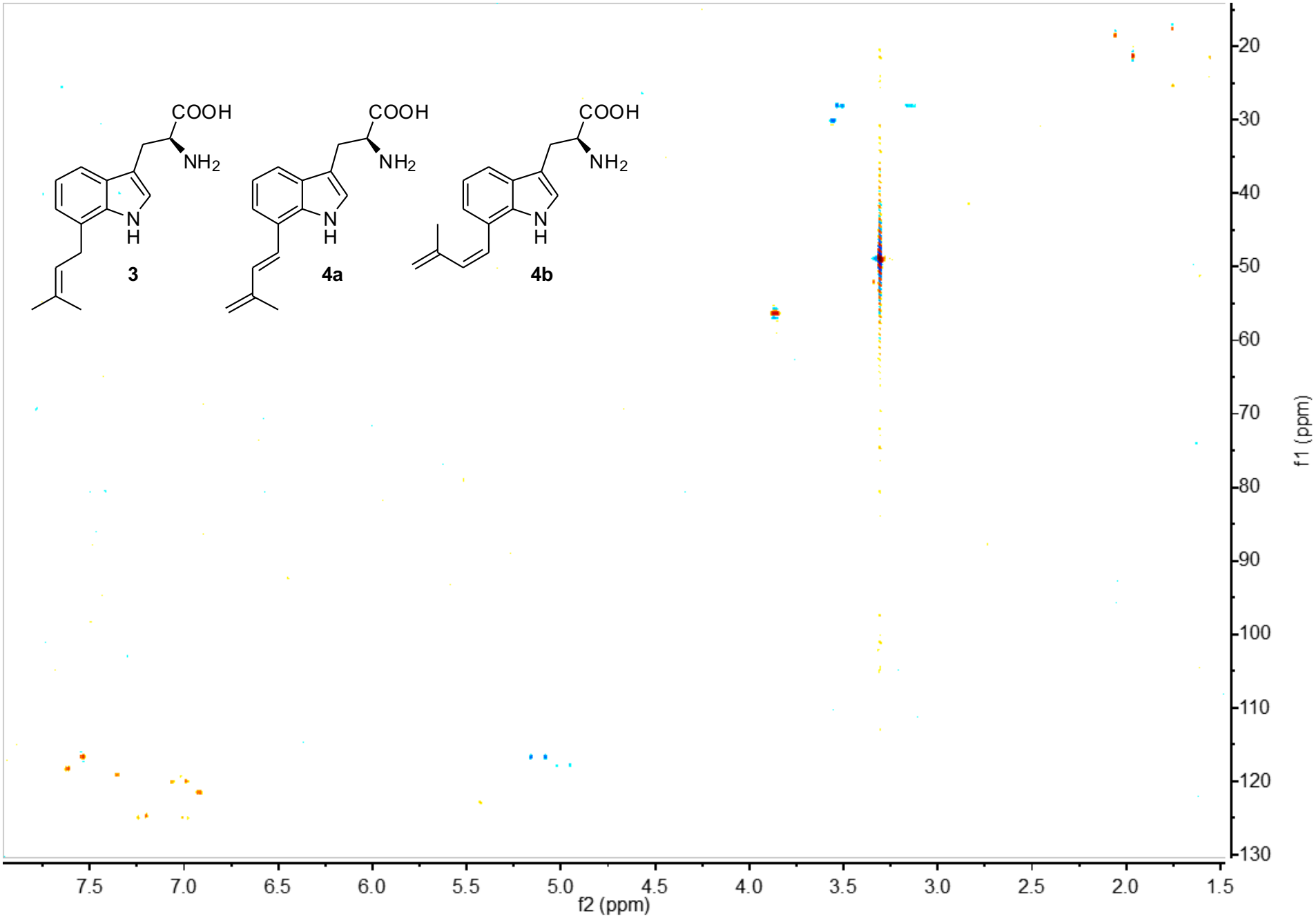
NMR spectra of compounds 3, 4a/b, 7, and 9



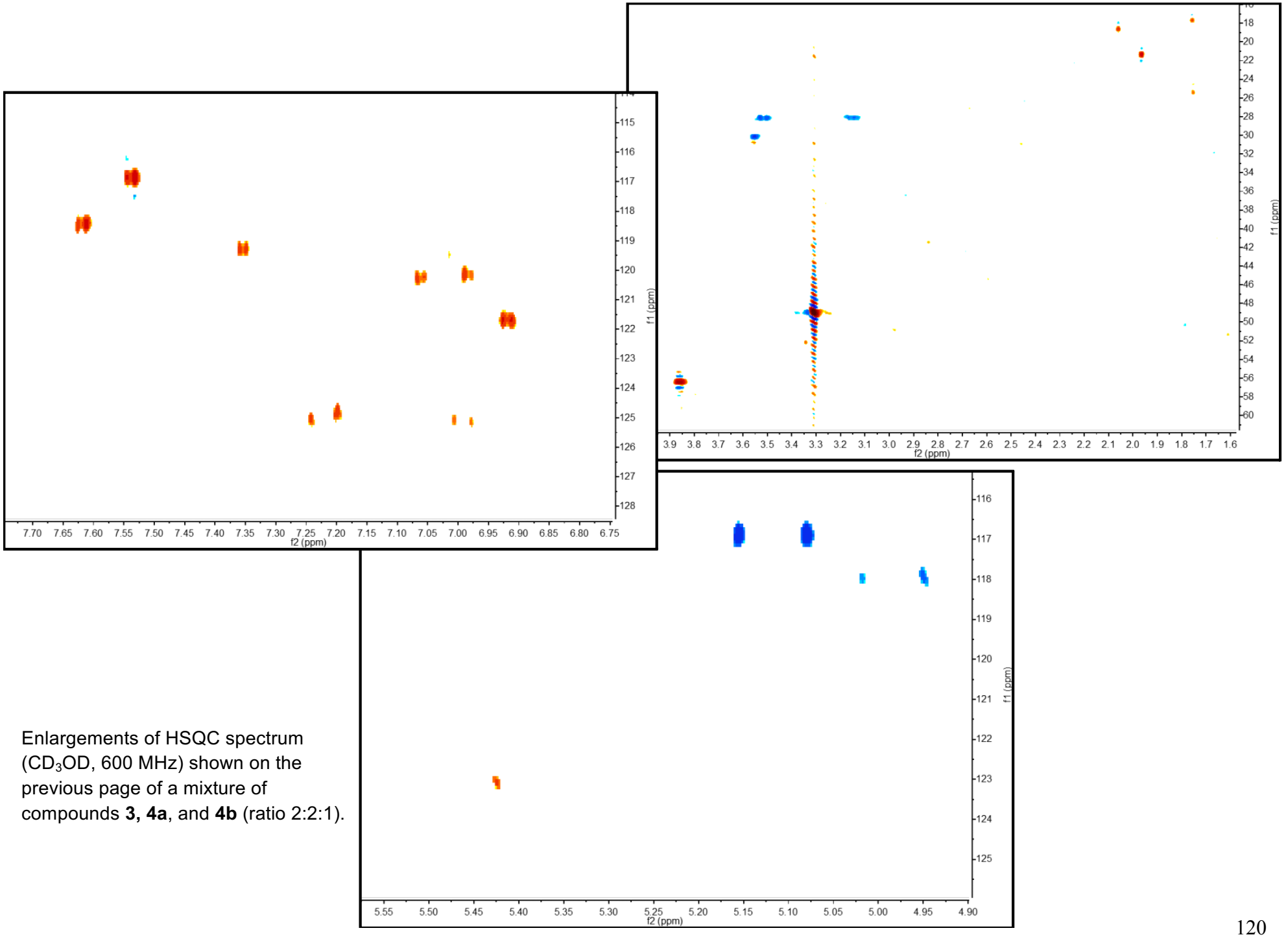
^{1}H NMR spectrum (CD_3OD, 600 MHz) of a mixture of compounds **3, 4a**, and **4b** (ratio 2:2:1).



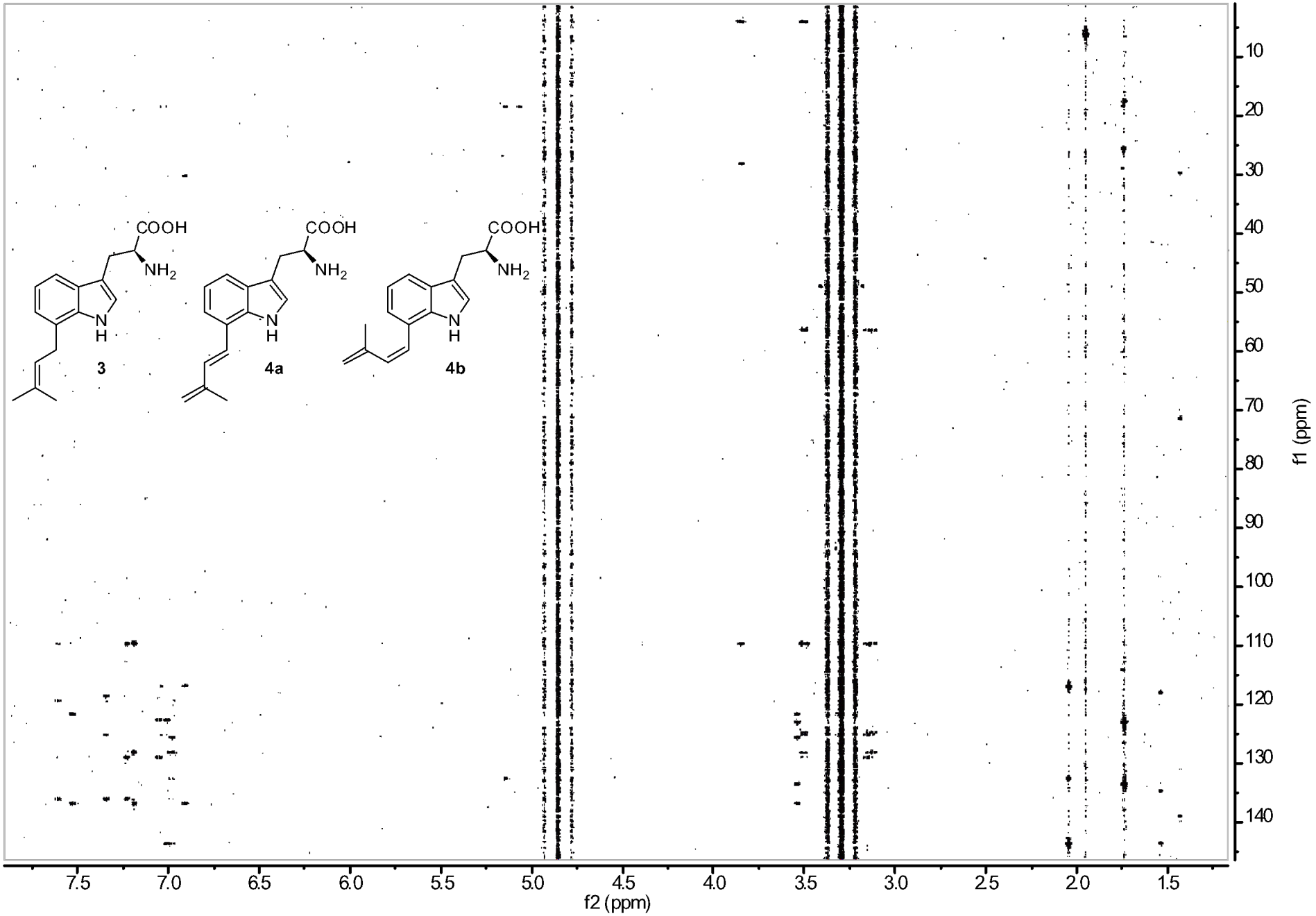
dqfCOSY spectrum (CD_3OD, 600 MHz) of a mixture of compounds **3, 4a**, and **4b** (ratio 2:2:1).



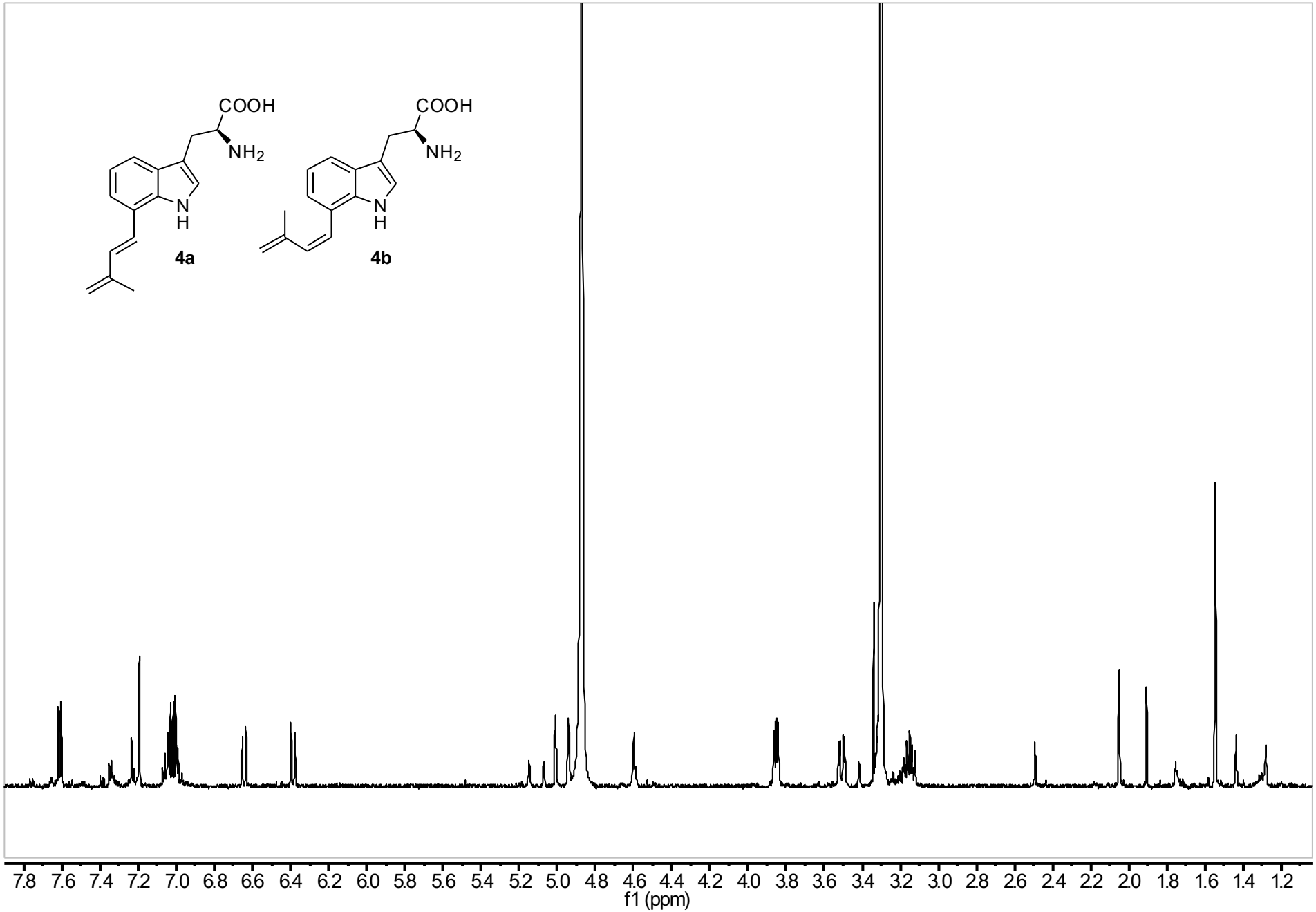
HSQC spectrum (CD_3OD, 600 MHz) of a mixture of compounds **3, 4a**, and **4b** (ratio 2:2:1).



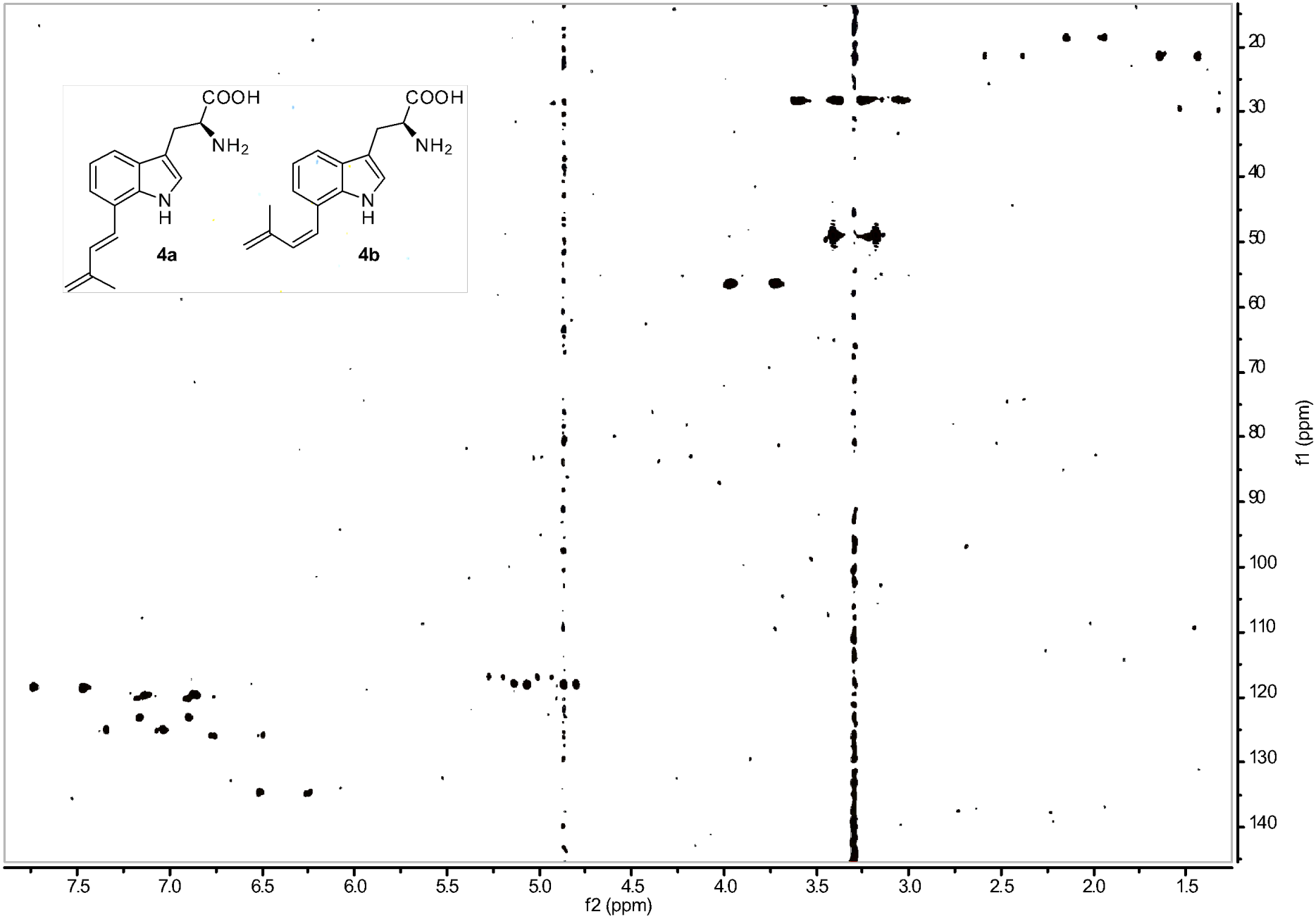
Enlargements of HSQC spectrum (CD_3OD, 600 MHz) shown on the previous page of a mixture of compounds **3, 4a**, and **4b** (ratio 2:2:1).



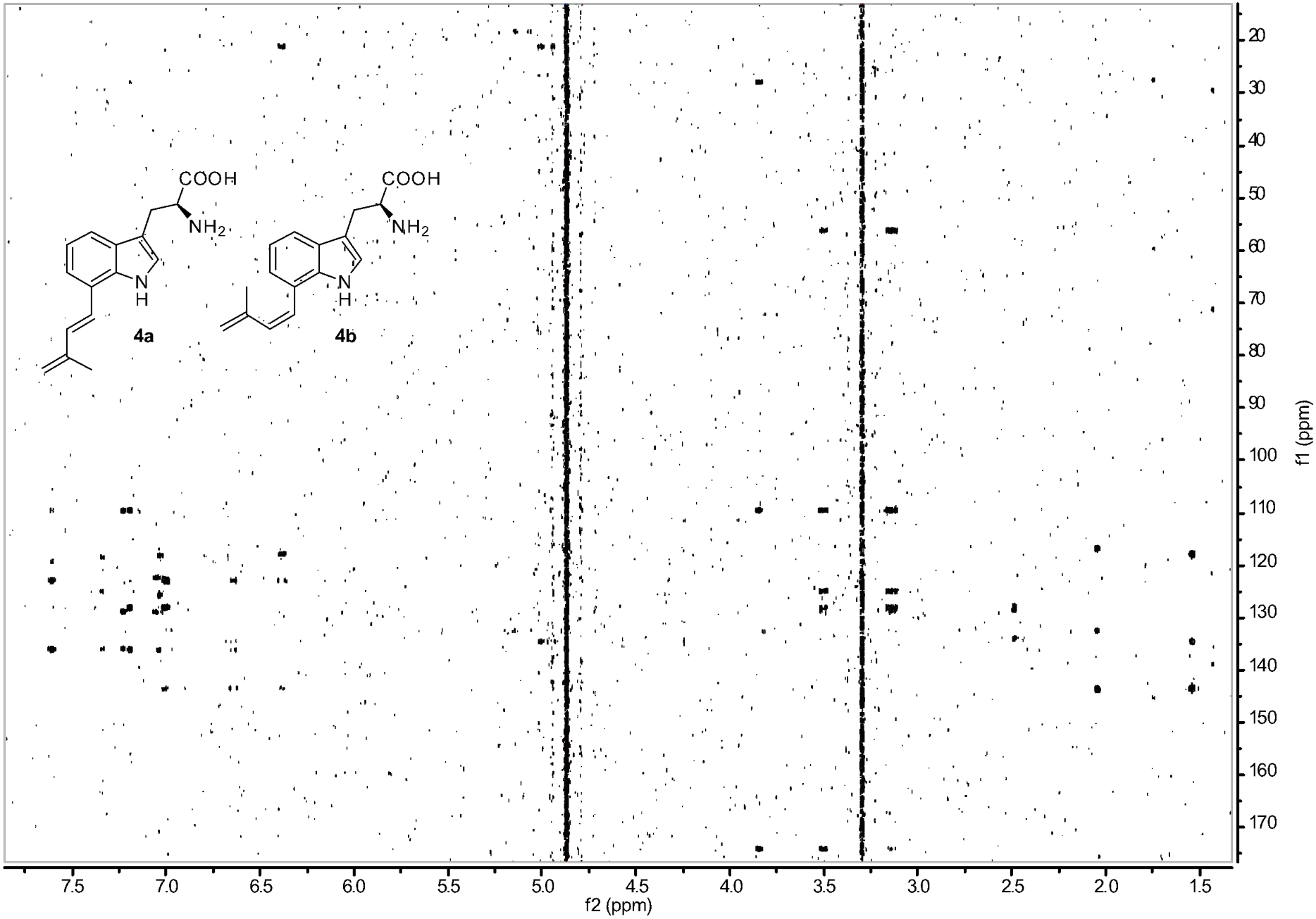
HMBC spectrum (CD_3OD, 600 MHz) of a mixture of compounds **3, 4a**, and **4b** (ratio 2:2:1).



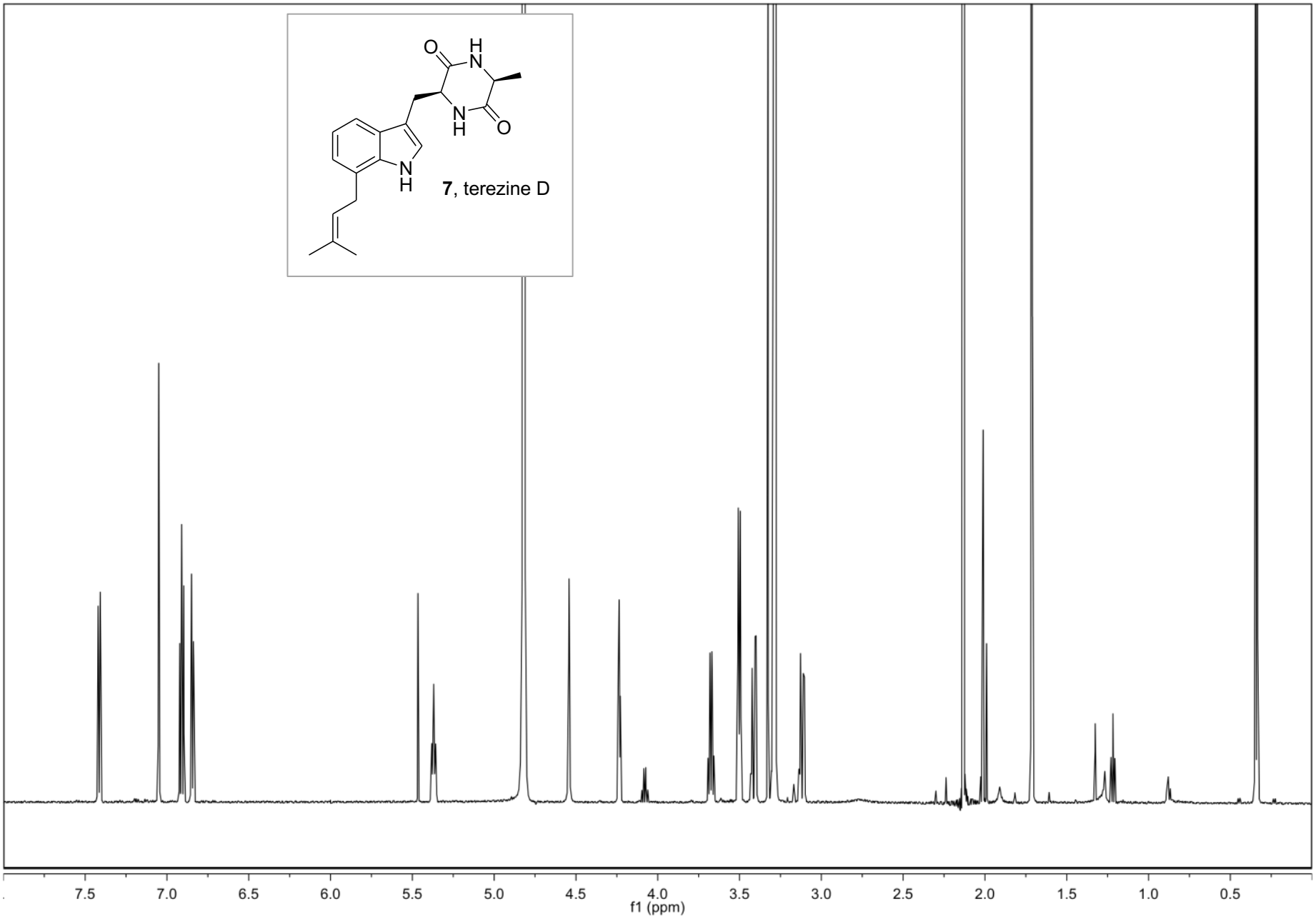
^{1}H NMR spectrum (CD_3OD, 600 MHz) of a mixture of compounds **4a** and **4b** (ratio 1:2.6).



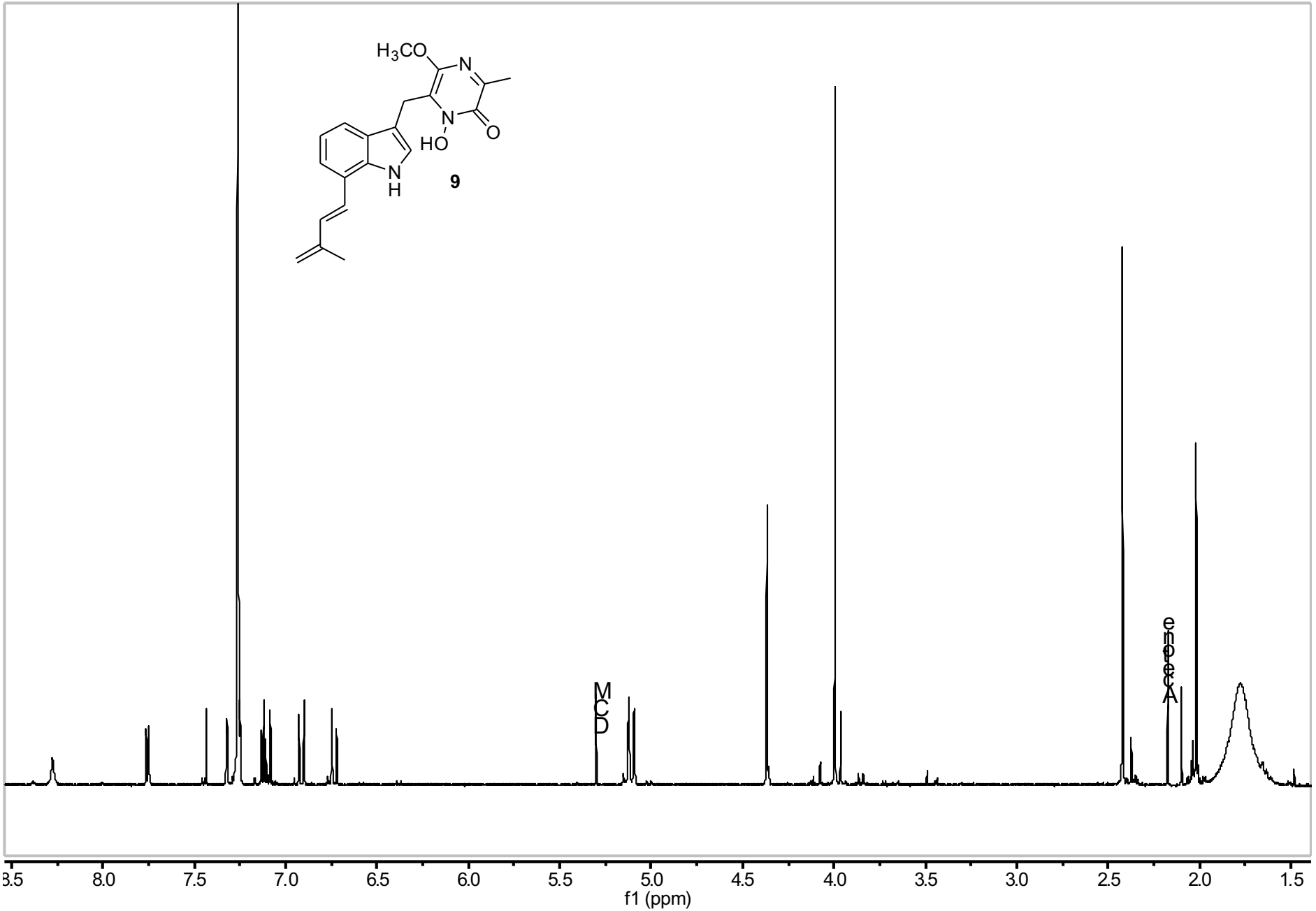
HMQC spectrum (CD_3OD, 600 MHz) of a mixture of compounds **4a** and **4b** (ratio 1:2.6).



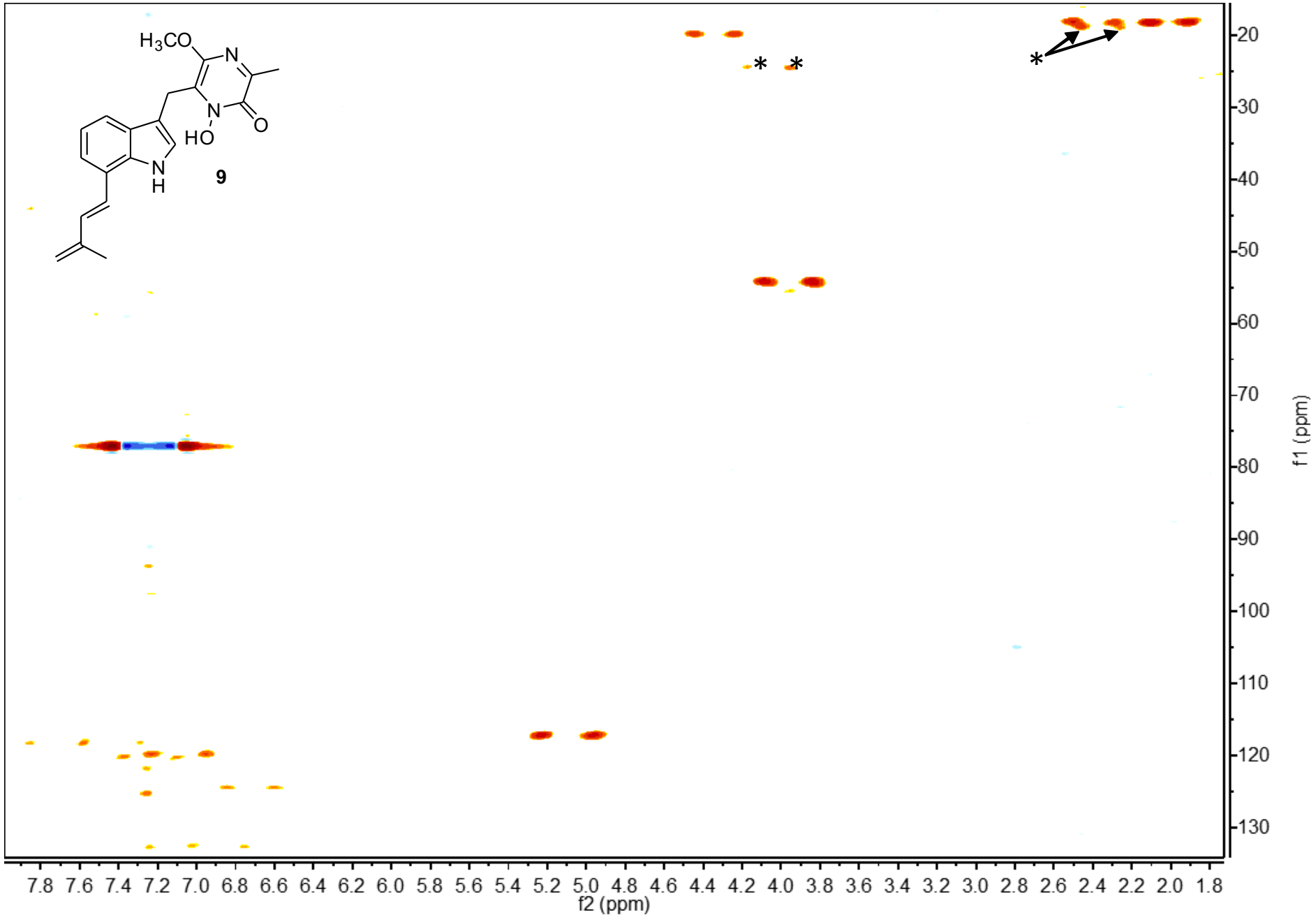
HMBC spectrum (CD_3OD, 600 MHz) of a mixture of compounds **4a** and **4b** (ratio 1:2.6).



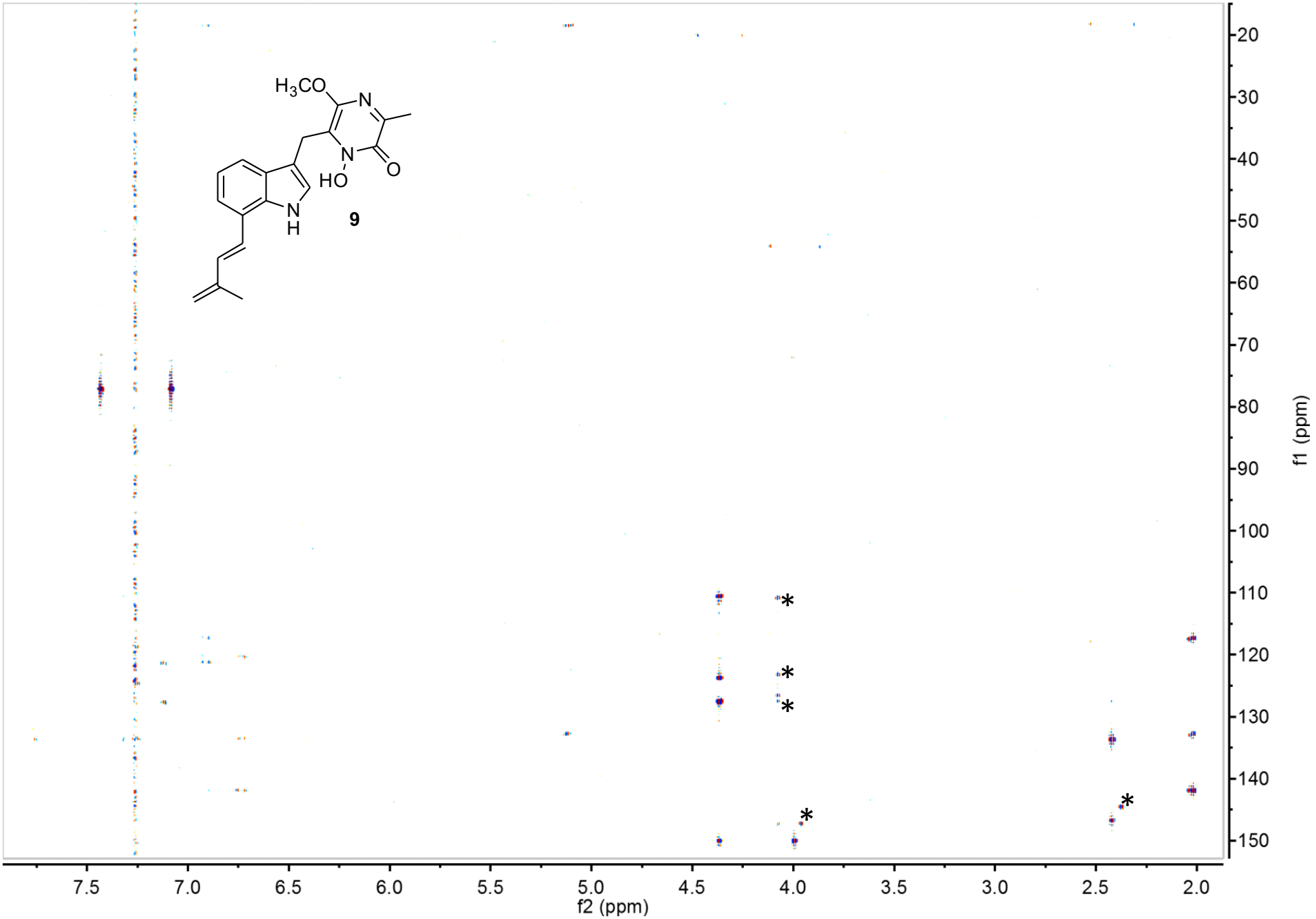
^{1}H NMR spectrum (CD_3OD, 600 MHz) of compound **7** (terezine D).



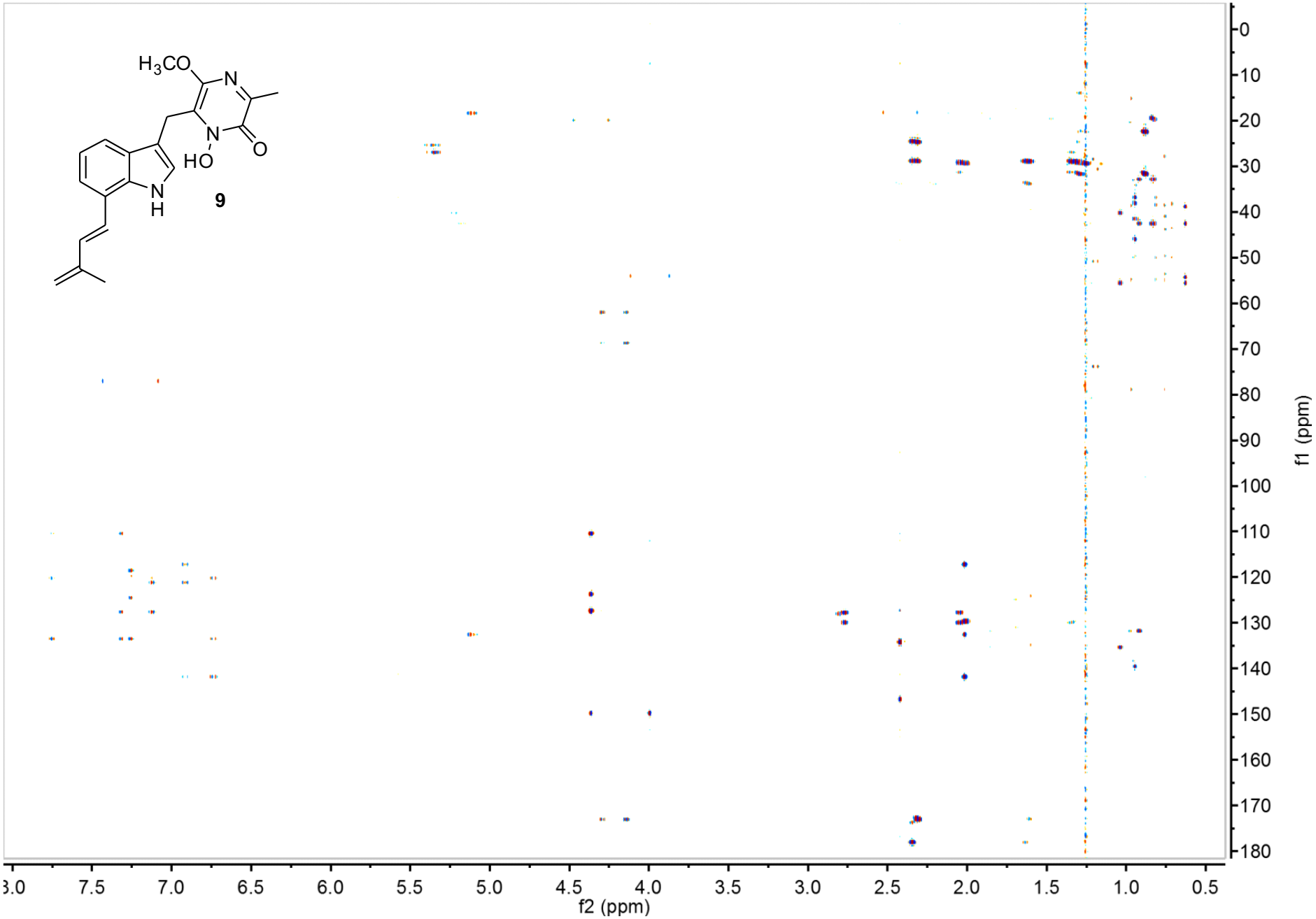
^{1}H NMR spectrum ($CDCl_3$, 600 MHz) of compound **9**.



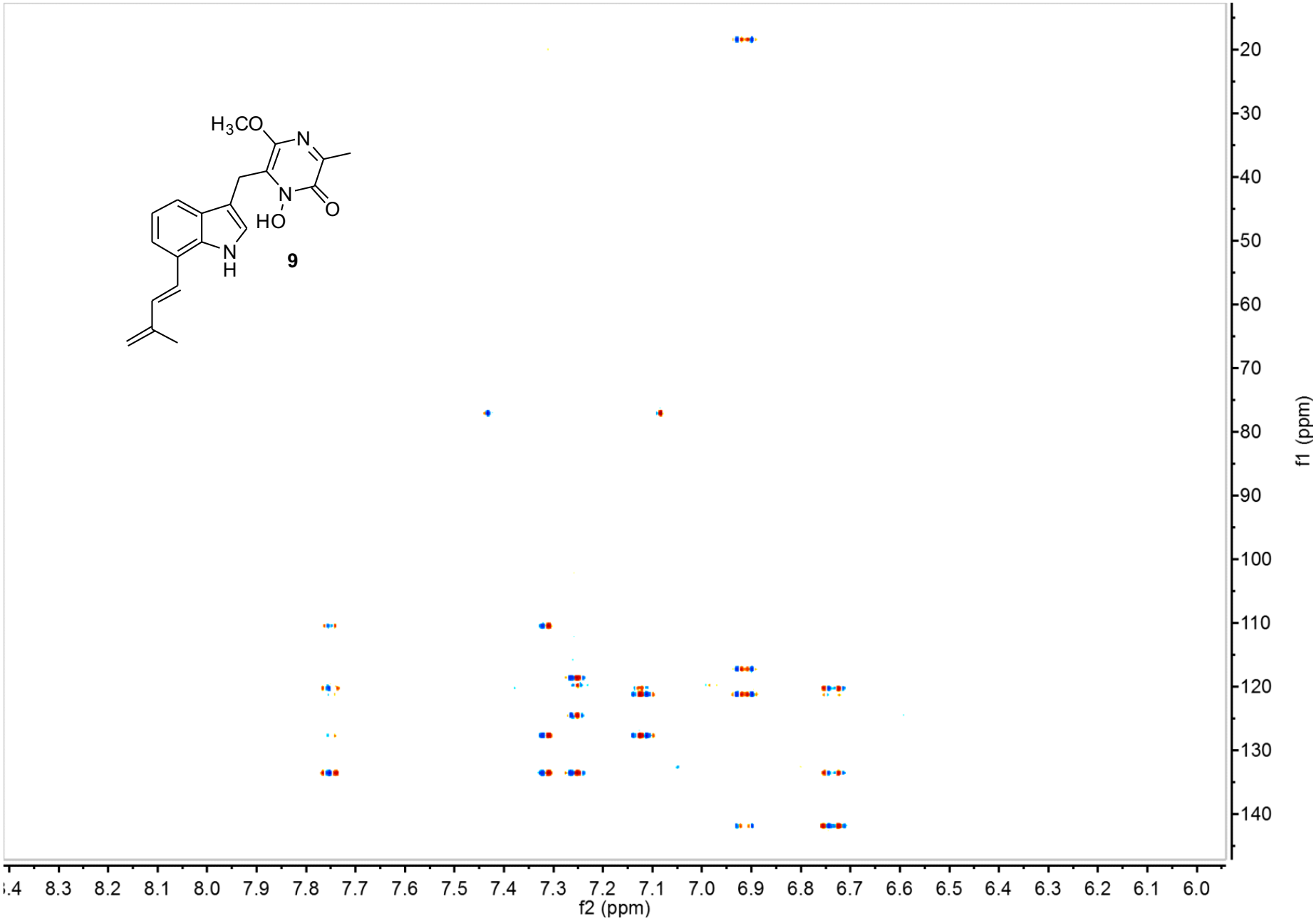
HMQC spectrum ($CDCl_3$, 600 MHz) of compound **9** (*: impurities resulting from partial decomposition during acquisition).



HMBC spectrum ($CDCl_3$, 600 MHz) of compound **9** (*: impurities resulting from partial decomposition during acquisition).



Aromatic region of HMBC spectrum ($CDCl_3$, 600 MHz) of partially purified sample of **9** (containing triglycerides).



Aromatic region of HMBC spectrum ($CDCl_3$, 600 MHz) of partially purified sample of **9** (containing triglycerides with signals below 6 ppm).

REFERENCES

1. Shimizu, K., Keller, N. P. Genetic involvement of a cAMP-dependent protein kinase ina G protein signaling pathway regulating morphological and chemical transitions in Aspergillus nidulans. *Genetics* **157**, 591-600, (2001).

2. Yu, J. H., Hamari, Z., Han, K. H., Seo, J. A., Reyes-Dominguez, Y., Scazzocchio, C. Double-joint PCR: A PCR-based molecular tool for gene manipulations in filamentous fungi. *Fungal Genet Biol* **41**, 973-981, (2004).

3. Yin, W. B., *et al*. An Aspergillus nidulans bZIP response pathway hardwired for defensive secondary metabolism operates through aflR. *Mol Microbiol* **83**, 1024-1034, (2012).

4. Bok, J.; Keller, N. P. *Fast and easy method for construction of plasmid vectors using modified Quick-change mutagenesis.*; Humana Press, 2012.

5. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular cloning : a laboratory manual*; 2nd ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., 1989.

6. Bok, J. W., *et al*. Chromatin-level regulation of biosynthetic gene clusters. *Nat Chem Biol* **5**, 462-464, (2009).

7. Arai, K.; Sato, S.; Shimizu, S.; Nitta, K.; Yamamoto, Y. Metabolic Products of Aspergillus terreus. VII. Astechrome: an Iron-containing Metabolite of the Strain IFO 6123. *Chem Pharm Bull* **29**, 1510-1517, (1981).

8. Xue, T.; Nguyen, C. K.; Romans, A.; Kontoyiannis, D. P.; May, G. S. Isogenic auxotrophic mutant strains in the Aspergillus fumigatus genome reference strain AF293. *Arch Microbiol* **182**, 346-353, (2004).

9. da Silva Ferreira, M. E., *et al*. The akuB^{KU80} Mutant Deficient for Nonhomologous End Joining Is a Powerful Tool for Analyzing Pathogenicity in Aspergillus fumigatus. *Eukaryot Cell* **5**, 207-211, (2006).

10. Lim, F.Y., *et al*. Genome-Based Cluster Deletion Reveals an Endocrocin Biosynthetic Pathway in Aspergillus fumigatus. *Appl. Environ Microbiol* **78**, 4117-4125, (2012).

11. Amaike, S., Keller, N. P. Distinct Roles for VeA and LaeA in Development and Pathogenesis of Aspergillus flavus. *Eukaryot Cell* **8**, 1051-1060 (2009).

12. Alexander, N. J., Hohn, T. M., McCormick, S. P. The TRI11 Gene of Fusarium sporotrichioides Encodes a Cytochrome P-450 Monooxygenase Required for C-15 Hydroxylation in trichothecene Biosynthesis. *Appl Environ Microbiol* **64**, 221-225, (1998).

13. Kremer, A., Westrich, L., Li, S. M. A 7-Dimethylallyltryptophan Synthase From Aspergillus Fumigatus: Overproduction, Purification and Biochemical Characterization. *Microbiology* **153**, 3409-3416, (2007).

14. Wang, Y.; Gloer, J. B.; Scott, J. A.; Malloch, D. Terezines A-D: new amino acid-derived bioactive metabolites from the coprophilous fungus sporormiella teretispora. *J Nat Prod* **58**, 93-99, (1995).

APPENDIX B

SECOND CONDENSATION-LIKE DOMAIN IN GLIP IS ESSENTIAL FOR DIKETOPIPERAZINE CYCLIZATION IN GLIOTOXIN BIOSYNTHESIS

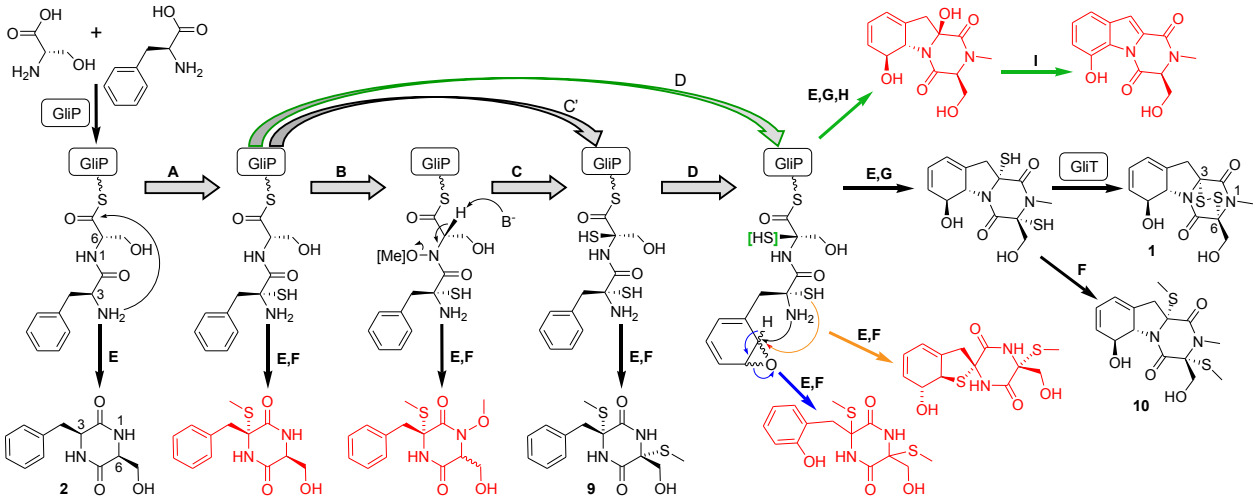


Figure B1. Model for gliotoxin biosynthesis, adapted from Forseth, *et al*[1].

```
LAEQVELVRRKRQASTAKPRTICDAFPDHCLSPLERQTWFQYLIAADVRTFNIPVLLHLGGTFDRDRLV
QSFNAVLASRKIFRTNFVETSLGPCRIFRDTPPRVLVCDGALDTTKEIDRSFDLARDELIRVFLDRRTL
LVVTSHAVADLNSVQNLLQDVSGVYAGRTTPTPDRWHYPRAPAWSRQATEQERKFWSKYLEGAPQRLDI
PRYPGQMAFEGRSRVSEFKGDLVRRAVTLGQEHGLSQHQLVCAAVAQTLQWLAGSNDVVLGSPWANRGH
TVEQESMGLFLDRLPLRFKTPVNADCATILQSTRAASQAAVCNSIPFEQVLNLLHLPRTIRQHPLFEAM
VTFHLKGAVEDCLAIEGLEVKREMCFASGAKFLLMFEWTEIEADHWTLRIEYDDHQLDDATVTTIEDSI
RCVLEGLADRLSRAAIHERLNAMHKTARTKVDWNFYRRLVGILQREMATCLGVSLDEFPCSVSFFEAGG
DSIQAWRLSRQLKRVGLEVPICNIFDHPTAQDLAQRLYRQVL
```

Figure B2. Amino acid sequence of C_2T_3 domains of GliP (Af293). Residues highlighted in yellow are conserved across C_T domains[2].

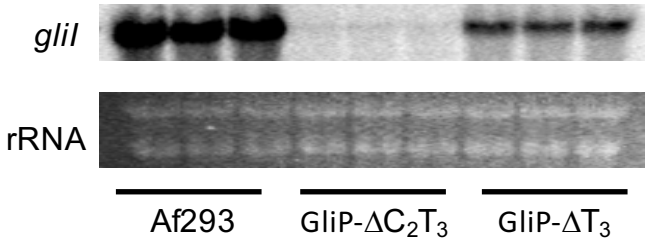


Figure B3. *gliI* gene expression in WT(Af293) GliP-ΔC_2T_3 and GliP-ΔT_3 A. fumigatus strains**.** 10^7 spores/mL were inoculated in 50mL liquid glucose minimal medium and incubated for 72hr at 25 ^oC, 225rpm, before total RNA extraction.

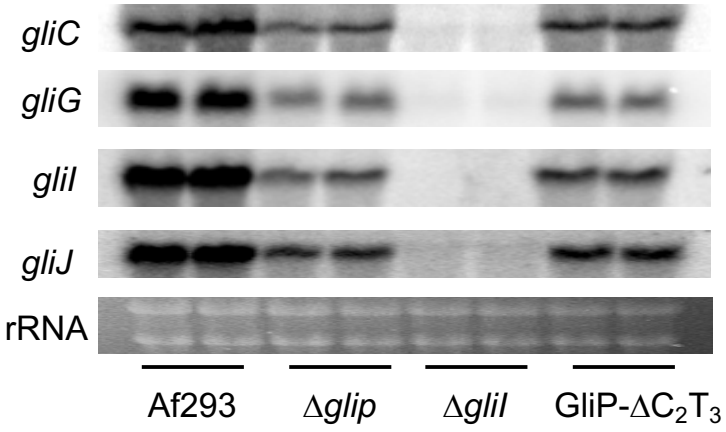


Figure B4. Gene expression of *gliC, gliG, gliI* and *gliJ* in WT(Af293), ΔGliP, ΔGliI, and GliP-ΔC_2T_3 A. fumigatus strains. 10^7 spores/ml were inoculated in 50ml liquid glucose minimal medium and incubated for 24hr at 30 ^oC, 225rpm and additional 24hr at 25^oC, 225rpm before adding gliotoxin at the concentration of 25mg/ml. After 24hr further cultivation at 25^oC, total RNA was isolated from samples for gene expression analysis.

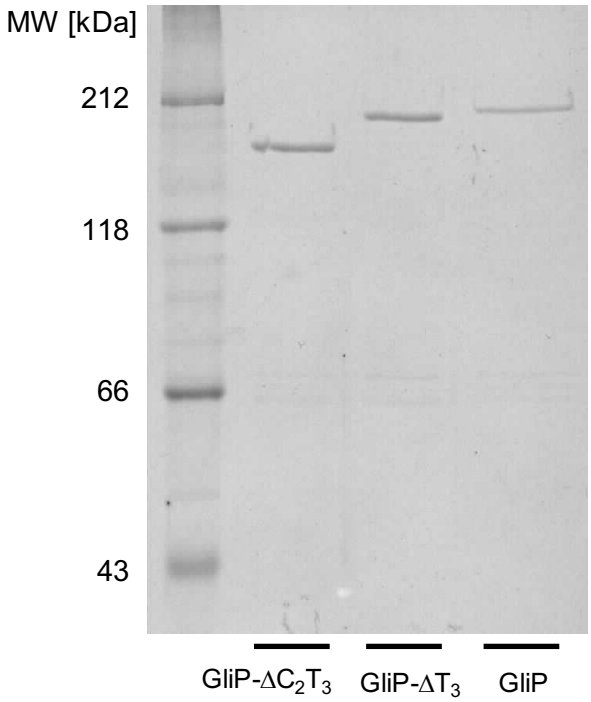


Figure B5. SDS-PAGE confirmation of recombinant GliP-ΔC_2T_3, GliP-ΔT_3, and native GliP.

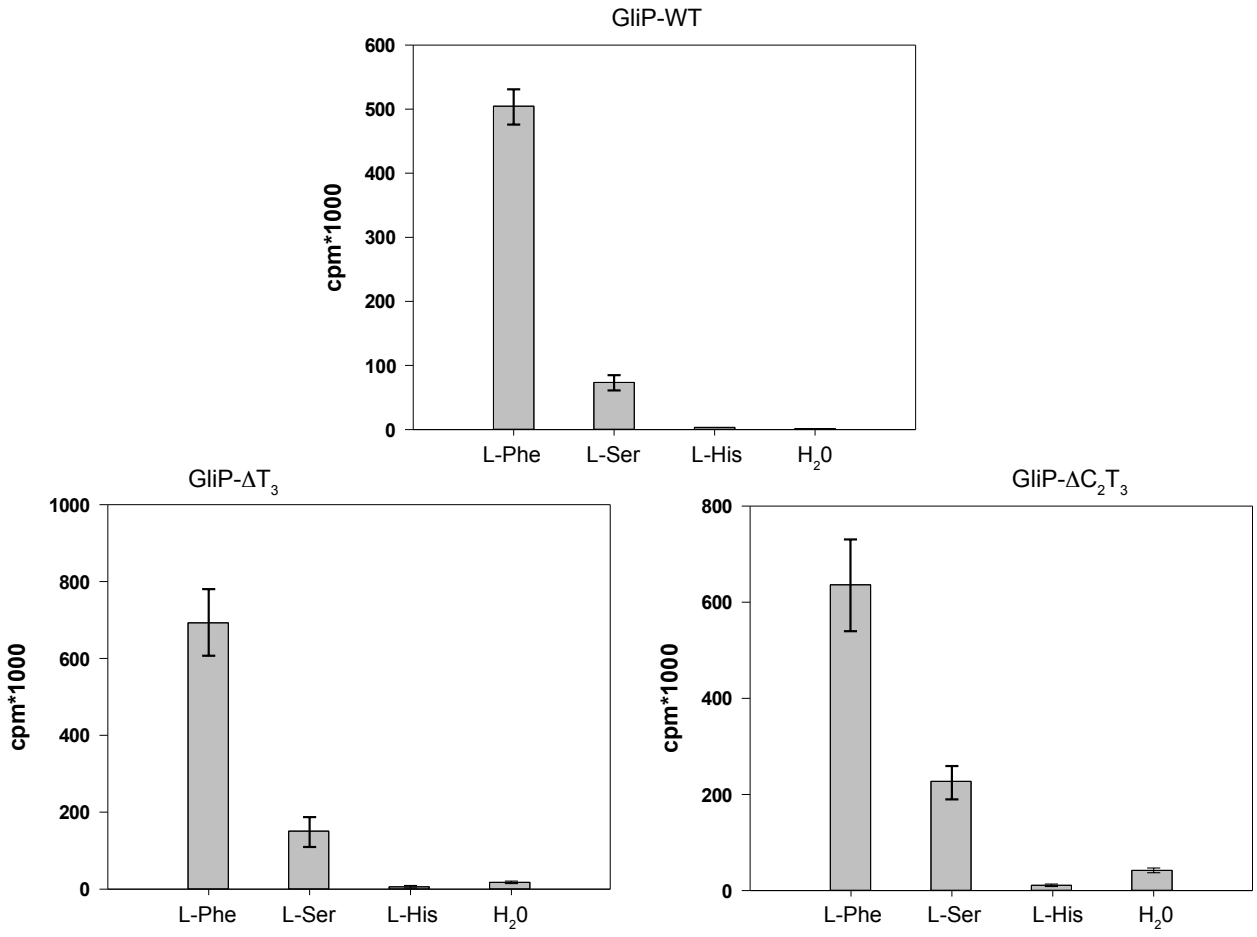


Figure B6. PPi-ATP exchange assay results for recombinant GliP-WT, GliP-ΔT_3, and GliP-ΔC_2T_3. Shown are raw turnover rates for L-Phe, L-Ser, L-His, and water for each enzyme[3], each assay was run in triplicate.

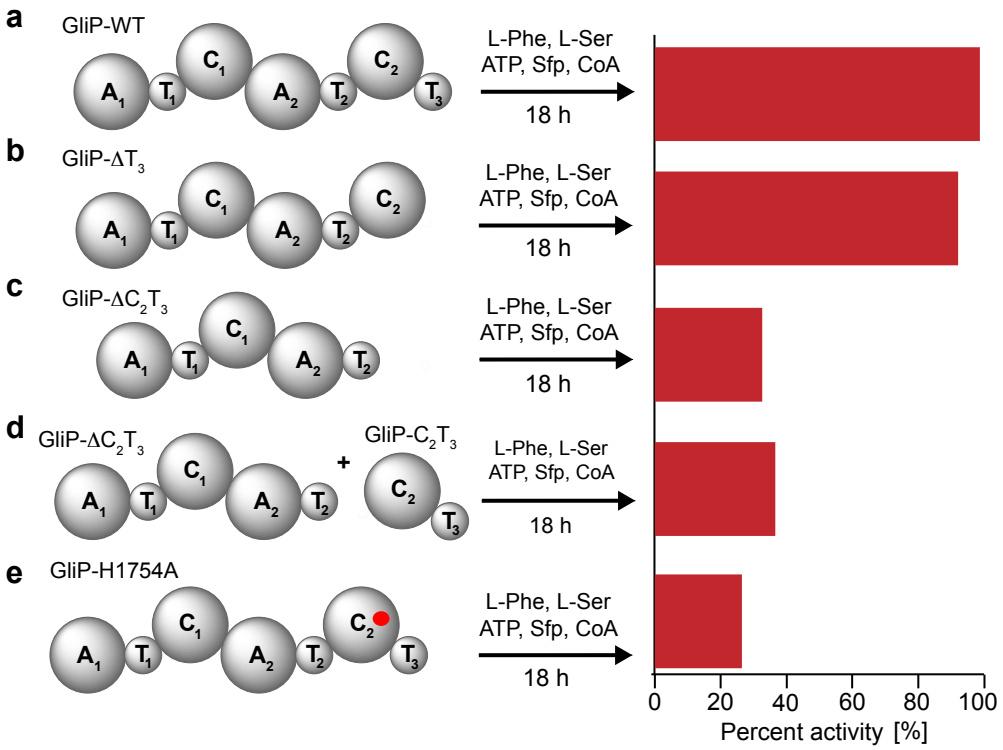


Figure B7. Relative activity of various GliP proteins for product formation *in vitro*. Percent activity measured as the amount of **2** and **12** produced by (**a**) GliP-WT, (**b**) GliP-ΔT_3, (**c**) GliP-ΔC_2T_3, (**d**) GliP-ΔC_2T_3 + GliP-C_2T_3 standalone, and (**e**) GliP-H1754A.

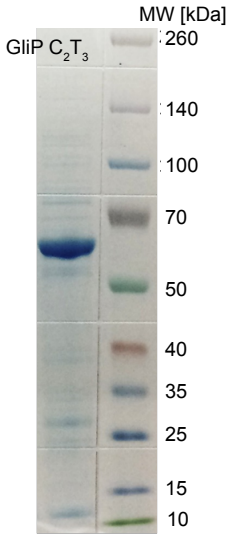


Figure B8. SDS-PAGE confirmation of recombinant GliP-C_2T_3.

GliP-C_2T_3 amino acid sequence

…….VSEFKGDLVRRAVTLGQEHGLSQHQLVCAAVAQTLQWLAGSNDVVLGSPWAN
RGHTVEQESMGLFLDRLPLRFKTPVNADCATILQSTRAASQAAVCNSIPFEQVLNLLH
LPRTIRQHPLFEAMVTFHLKGAVEDCLAIEGLEVKREMCFASGAKFLLMFEWTEIEAD
HWTLRIEYDDHQLDDATVTTIEDSIRCVLEGLADRLSRAAIHERLNAMHKTARTKVD
WNFYRRLVGILQREMATCLGVSLDEFPCSVSFFEAGGDSIQAWRLSRQLKRVGLEVPI
CNIFDHPTAQDLAQRLYRQVL*

[M+2H]$^+$ = 715.79

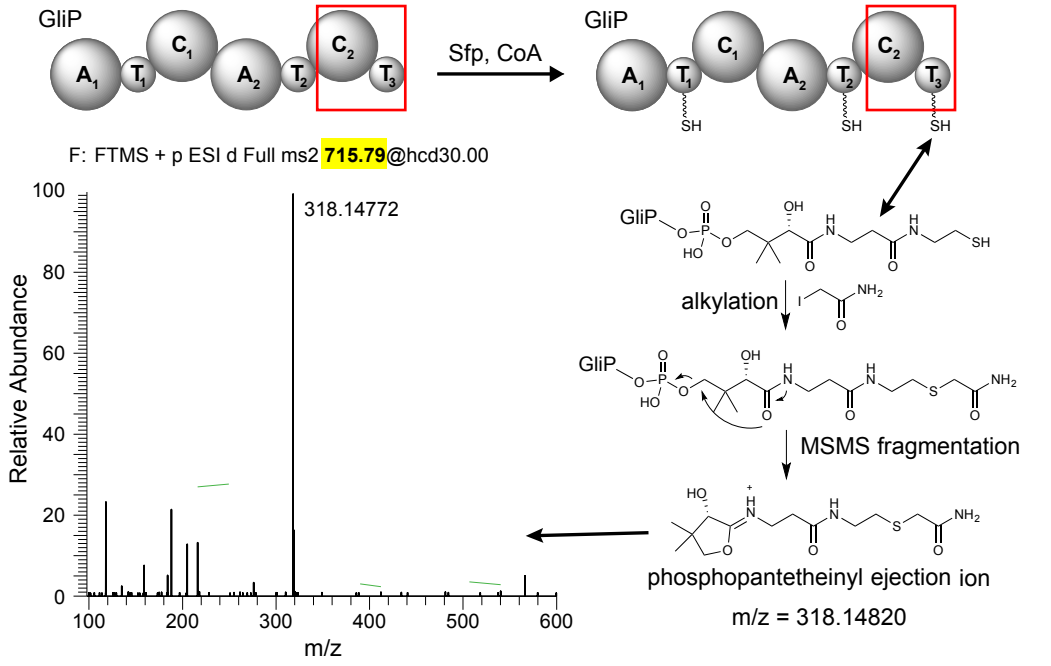


Figure B9. Mass spectrometric confirmation of posphopantetheinyl modification of GliP-T_3. See methods for experimental details.

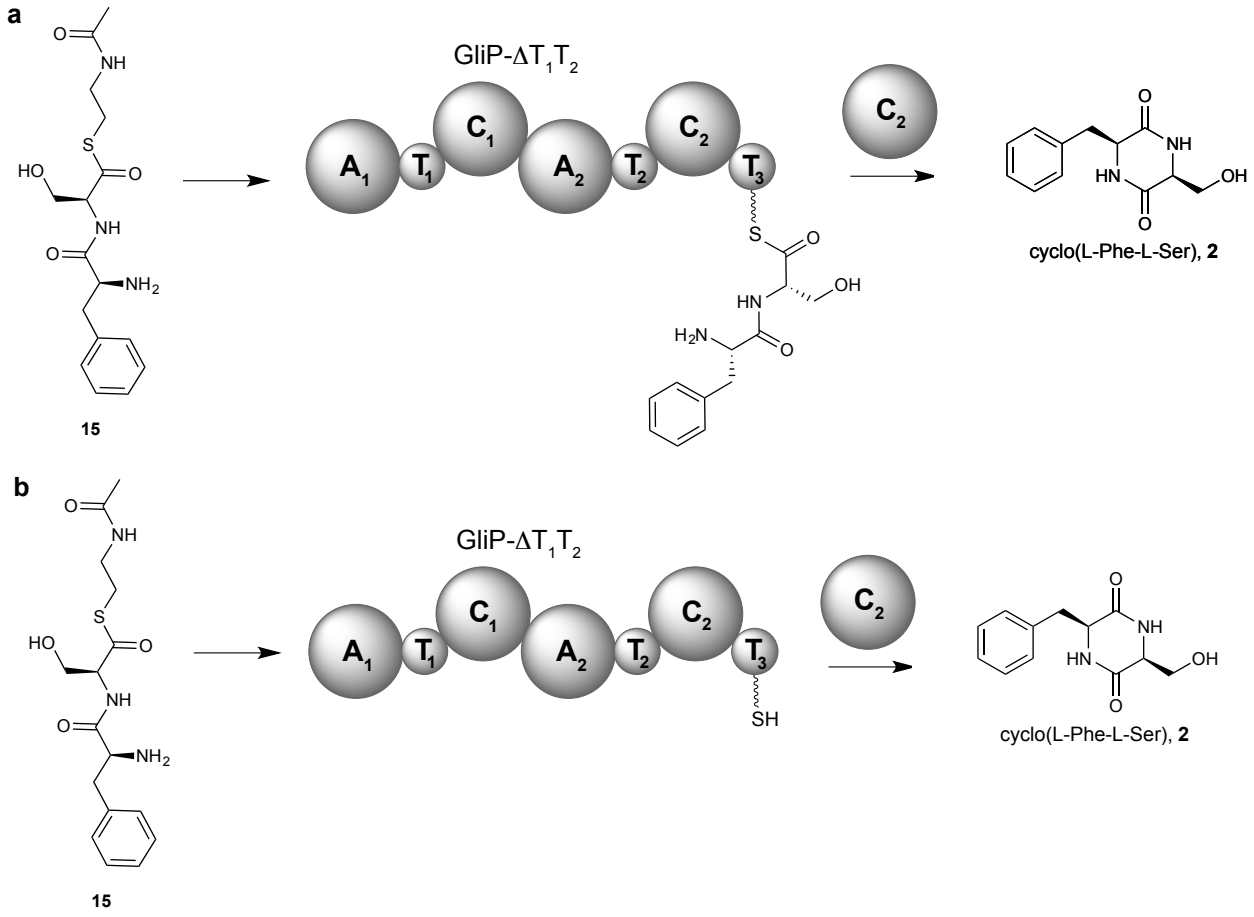


Figure B10. *In vitro* cyclization activity of GliP-ΔT_1T_2 toward **15.** (**a**) **15** can first be loaded onto T_3 via trans-thiolation, then cyclized by the C_2 domain to form **2**, or (**b**) the C_2 domain can directly cyclize **15** to form **2**.

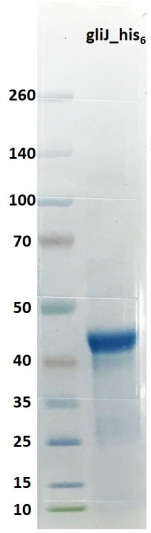


Figure B11. SDS-PAGE confirmation of H_6-tagged GliJ.

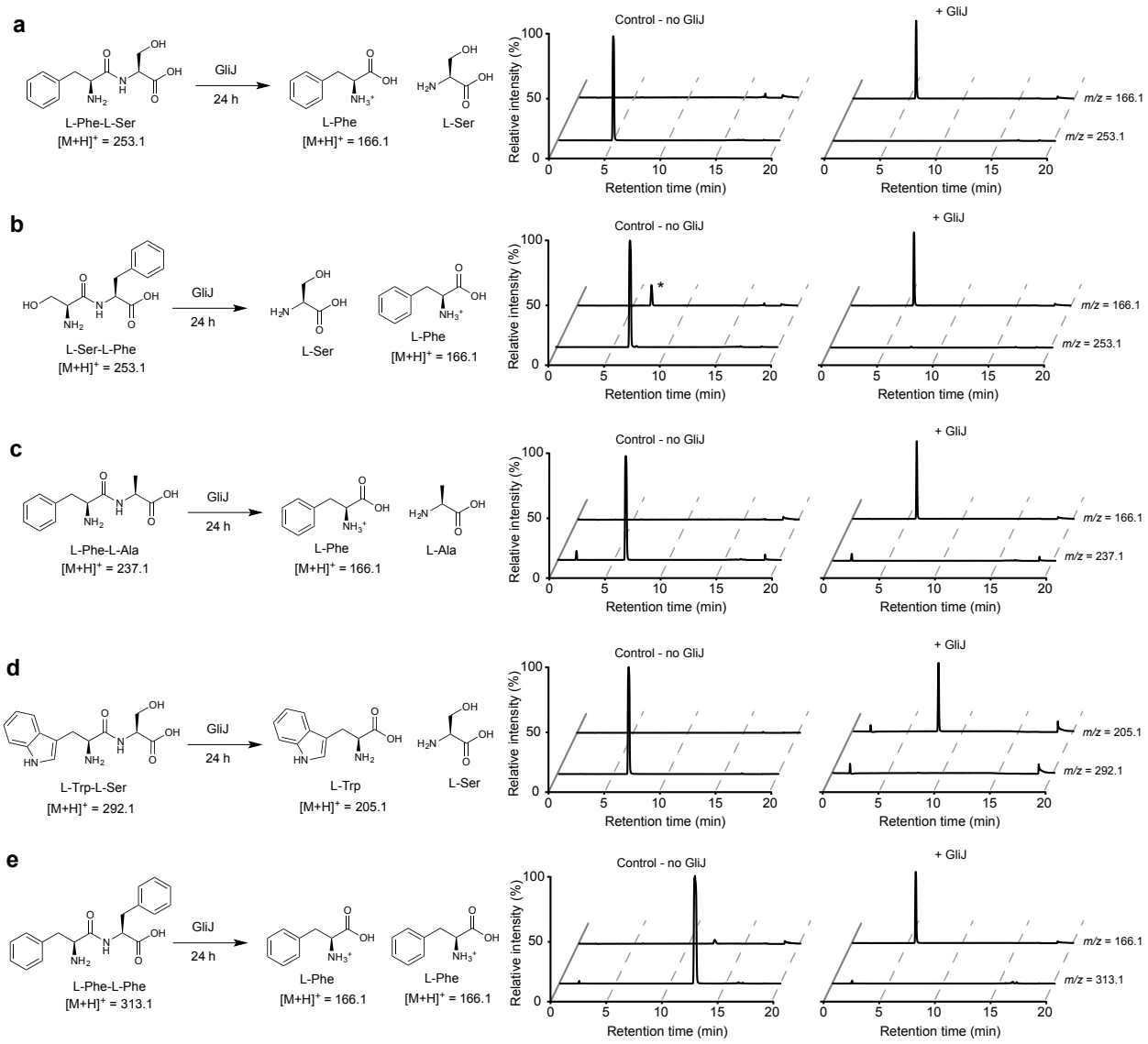


Figure B12. GliJ dipeptidase activity exhibits loose substrate specificity. (**a-e**) *In vitro* assay with recombinant GliJ and the indicated dipeptides showing LC-MS analysis with extraction-ion-chromatograms for dipeptides and the corresponding hydrolyzed amino acids in the control experiment (left set) and with addition of GliJ (right set).

Strains, media, and growth conditions. The fungal strains used in this study are listed in Supplementary Table 1. All strains were grown at 30 °C on glucose minimal medium (GMM[4]) and, when appropriate, were supplemented with 0.56 g uracil L^{-1}, 1.26 g uridine L^{-1}, 1.0 g arginine L^{-1} and maintained as glycerol stocks at -80 ºC. *Escherichia coli* strain DH5α was propagated in LB medium with appropriate antibiotics for plasmid DNA.

Gene cloning, plasmid construction, and genetic manipulations. To introduce the histidine to alanine amino acid substation in a strain of *Aspergillus fumigatus*, we first fully deleted *gliP* in a strain with *akuA* deleted (TFYL44.1) to increase the rate of homologous recombination and decrease the amount of transformants screened to obtain desired strain. The open reading frame of *gliP* was replaced with a copy of *pyrG* from *Aspergillus parasiticus* to complement the *pyrG* auxotrophy. To generate a construct to delete *gliP* the flanking regions of the *gliP* open reading frame were amplified (gliP3'-F & gliP3' R, and gliP5'-F & gliP5'-R) as well as the *A. parasiticus pyrG* gene (Ap-pyrGF & Ap-pyrGR). These PCR products were fused using double joint PCR and used to transform TFYL44.1 to create strain TBTP12.02 which was confirmed by Southern blot analysis[5].

Two plasmids were then generated, one which included a full length copy of *gliP* (pBTP12), and one that contained a H1754A copy of *gliP* (pBTP13), both targeted to the *akuA* locus. The pBTP12 plasmid was assembled by amplifying *akuA* flanks (KU5'-F & KU5'-R, and KU3'-F & KU 3'-R), *gliP* (gliP-F & gliP-R), and *A. fumigatus argB* as the selectable marker (AFU argB fwd & AFU argB rev). These PCR fragments were combined with digested pJMP218 to act as the plasmid backbone in a yeast transformation to allow for homologous recombination to assemble the fragments into a full plasmid. pBTP13 was assembled using the same fragments and method, expect the H1754A substitution

was introduced by using primers containing the mutation and amplifying *gliP* in two separate PCR reactions (gliP-F & gliP-H1754A-R, and gliP-H1754A-F & gliP-R). TBTP12.02 was then transformed with pBTP12 and pBTP13 to generate TBTP99 and TBTP100 respectively, which were confirmed by Southern blot analysis. TBTP12.02 was taken to prototrophy by amplifying *argB* (AFU argB fwd &AFU argB rev2) from *A. fumigatus* and selecting prototrophic transformants generating TBTP94.

Nucleic acid analysis. Plasmid preparation, digestion with restriction enzyme, gel electrophoresis, blotting, hybridization, and probe preparation were performed by standard methods[6]. *Aspergillus* DNA for diagnostic PCR was isolated using the previously described method[7]. Sequence data were analyzed using the LASERGENE software package from DNASTAR.

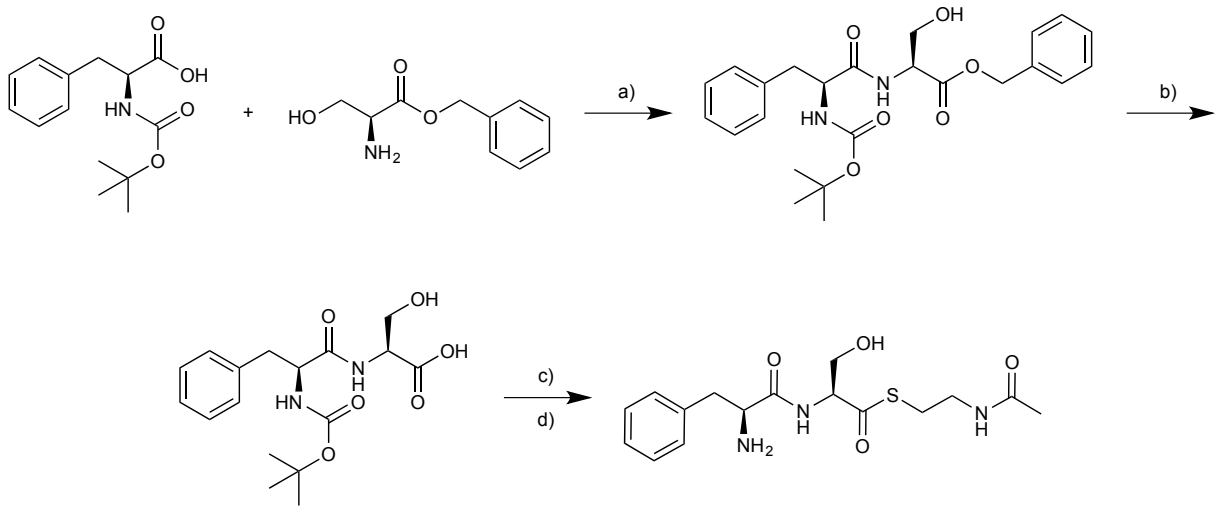
Northern analysis. 50 mL of liquid GMM were inoculated with 10^6 or 10^7 spores per mL of all appropriate strains in this study and incubated with shaking at 250 rpm at 25 °C. After 44 h, the mycelium was collected and total RNA was extracted by using Isol-RNA Lysis Reagent according to the manufacturer's instructions (5 Prime, Hamburg, Germany).

Fermentation and metabolome extraction. *A*. *fumigatus* strains were inoculated (1.0×10^6 spores/mL) into 25 mL GMM in a 50 mL flask at 30 °C with shaking at 220 rpm. After 5 days, liquid fungal cultures including fungal tissue and media were frozen using a dry ice acetone bath, and lyophilized. The lyophilized residues were extracted with 12.5 mL of mixture of acetonitrile, ethyl acetate, and water (80:15:5) for 0.5 h with vigorously

stirring. Extracts were filtered over cotton, evaporated to dryness, and stored in 8 mL vials. (a) Preparation for LC-HRMS Analysis: Crude extracts were suspended in 1.5 mL of extraction solvent and centrifuged to remove insoluble materials, and the supernatant was subjected to LC-HRMS analysis.

Analytical methods and equipment overview. (a) NMR Spectroscopy: NMR spectroscopic instrumentation: a Varian INOVA 600 MHz NMR spectrometer (600 MHz ^{1}H reference frequency, 151 MHz for ^{13}C) equipped with an HCN indirect-detection probe. Non-gradient phase-cycled dqfCOSY spectra were acquired using the following parameters: 0.6 s acquisition time, 400-600 complex increments, 16-32 scans per increment. Non-gradient HSQC, HMQC, and HMBC spectra were acquired with these parameters: 0.25 s acquisition time, 200-500 increments, 8-32 scans per increment. ^{1}H, ^{13}C-HMBC spectra were optimized for $J_{H,C}$ = 6 Hz. Susceptibility-matched NMR tubes (Shigemi) were used for sample amounts smaller than 1 mg. NMR spectra were processed and baseline corrected using Varian VNMR and MestreLabs MestReC and MNOVA software packages. (b) Mass spectrometry: High-resolution mass spectrometry was performed on a Waters nanoACQUITY UPLC system equipped with a Waters Acquity UPLC HSS C-18 column (2.1 x 100 mm, 1.8 µm particle diameter) connected to a Xevo G2 QTof Mass Spectrometer operated in electrospray positive (ESI^{+}) ionization mode. Low-resolution mass spectrometer was performed on an HPLC-MS system equipped with a diode array detector and connected to a Quattro II spectrometer (Micromass/Waters) operated in positive electrospray ionization (ESI^{+}) mode. Data acquisition and processing for the HPLC-MS was controlled by Water MassLynx software.

(c) Chromatography: Flash chromatography was performed using a Teledyne ISCO CombiFlash system. For semi-preparative HPLC a Supelco Discovery HS C-18 column (25 cm x 10 mm, 5 µm particle diameter) was used.



Synthesis of *N*-acetylcystamine-L-Phe-L-Ser (15). a) EDC, HOBt, DMF. b) H_2, Pd/C (10 mol %). c) *N*-acetylcystamine, EDC, HOBt, DMF. d) 40 % TFA, DCM[8]. See below for NMR assignments and spectra of **15**.

Synthesis of cyclo-L-Phe-L-Ser (2). In order to access **2** we simply allowed **15** to cyclize in 1 M phosphate buffer, pH 8 and then isolated **2** by semi-preparative HPLC, described above. See below for ^{1}H NMR of **2**[1].

Table B1. Fungal strains and plasmids used in this study

Strain/plasmid	Description
***Aspergillus fumigatus* 293 background strains**	

Af293	Wild type
Af293.1	*A. fumigatus pyrG1*
Af293.6	*A. fumigatus pyrG1, argB1*
ARC2	*ΔgliP::para pyrG1*
TJW139.3	*ΔC_2T_3 gliP:: para pyrG1*
TJW140.16	*ΔT_3 gliP:: para pyrG1*
TBTP94.1	pyrG1; ΔGliP::A.p pyrg1; KO:KU70
TBTP99.1	pyrG1; ΔGliP::A.p pyrg1; GliP::Af. argB@KU70
TBTP100.6	pyrG1; ΔGliP::A.p pyrg1; GliP-H1754A::Af. argB@KU70

Plasmids

Name	**Sequence**	**Purpose**
gliP 3'-F	GATAGCACACCCTCGGAATAGTCCTCTCGGCGTTCCATTCGACAGAAGACGAGG	Deletion
gliP 3'-R	CTGGAGCAGCTTCCGTGC	Deletion
gliP 5'-F	GAGGCTCTGCTCAGATGAGG	Deletion
gliP 5'-R	CGATGATAAGCTGTCAAACATGAGGCAGAGCGTAGGGTTGAGC	Deletion
Ap-pyrG-F	CTCATGTTTGACAGCTTATCATCG	Deletion
Ap-pyrG-R	CCGAGAGGACTATTCCGAGG	Deletion
KU5'-F	TATTGCCGTTGGATCTTTGGGG	Complement
KU5'-R	GGTATGGATTGTCATCAGCCATAGTGAG	Complement
gliP-F	TTCTCACTATGGCTGATGACAATCCATACCCGCTCGCCAATATGCTTGC	Complement
gliP-R	GAAAATTTGTCTTGGATGCAGACCGCGTTCCTGTGACGAACTCGACGAGG	Complement
gliP-H1753A-R	ACGGCTGCGCTGGTAACCACAAGGAGG	Complement Mutant Copy
gliP-H1753A-F	GCACCCTCCTTGTGGTTACCAGCGCAGCCGTCGCCGATCTCAACAGCGTG	Complement Mutant Copy
AFU argB fwd	GAACGCGGTCTGCATCCAAG	Complement

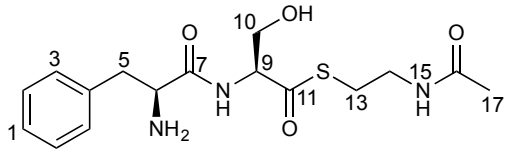
AFU argB rev	TGGTTAGTAACATTCAGACAGTCGGCATGCAGGGACTGAACCTGGTGAATCG	Complement
KU3'-F	GCATGCCGACTGTCTGAATGTTACTAACC	Complement
KU3'-R	TCACATGTTCTTTCCTGCGTTATCCCCTACACCAAGAAGCTCACCACCCC	Complement
AFU argB rev2	AGCATCCATTCTGCGTCTCG	Complement

Table B2. LC-HRMS data of reported compounds

Compound	HR-ESI+ Observed (m/z)	Ion	Calculated Ion Formula	Calculated m/z	Retention time [min]
1	263.1030	$[M\text{-}S_2+H]^+$	$C_{13}H_{15}N_2O_4^+$	263.1032	5.00
2	235.1080	$[M + H]^+$	$C_{12}H_{15}N_2O_3^+$	235.1082	2.30
9	279.0801	$[M\text{-}SCH_3+H]^+$	$C_{13}H_{16}N_2O_3S^+$	279.0803	3.84
10	309.0906	$[M + H]^+$	$C_{14}H_{18}N_2O_4S^+$	309.0908	5.00

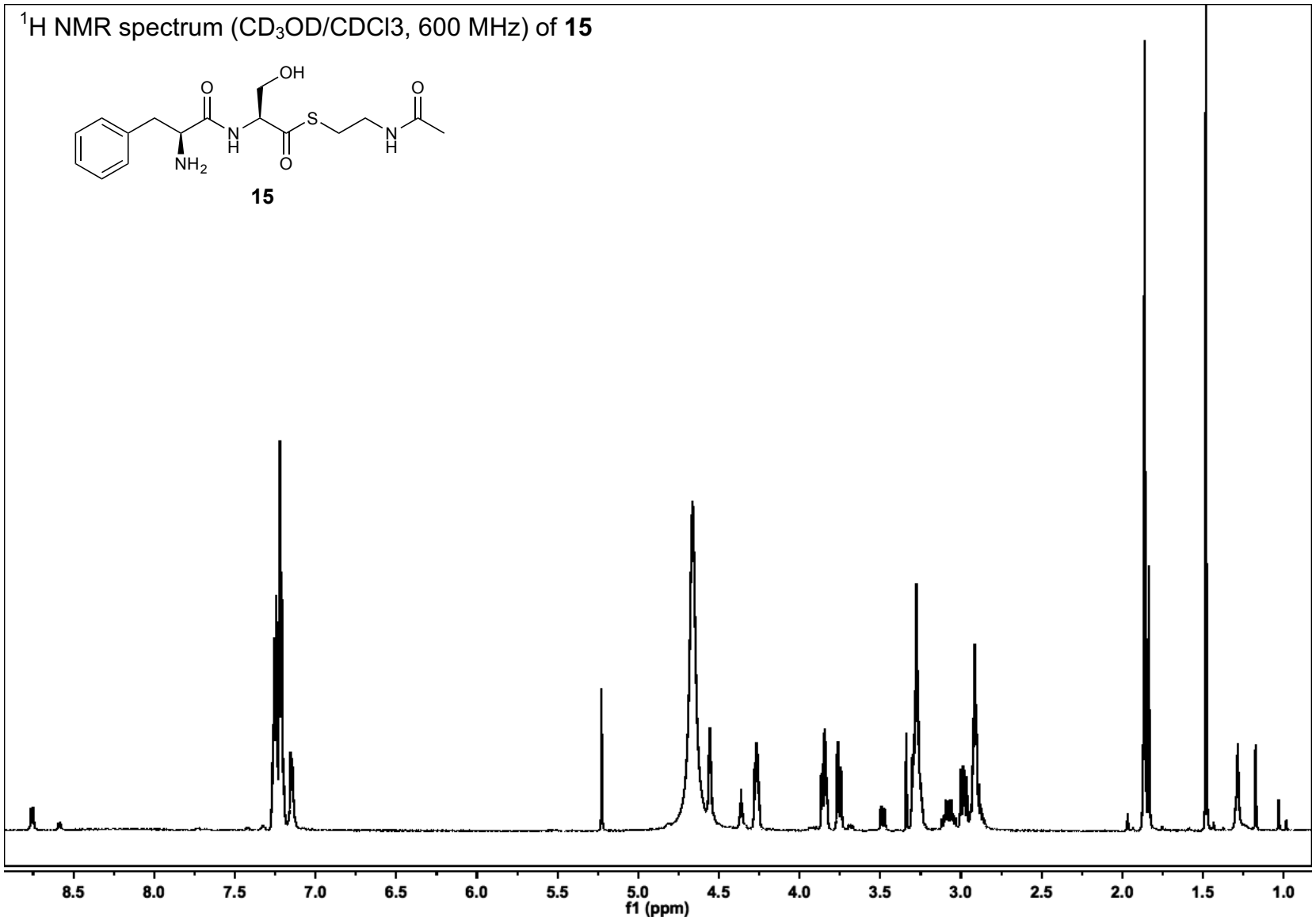
Table B3. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for compound **15** in a 80:20 mixture of methanol-d_4 : chloroform-d_3.

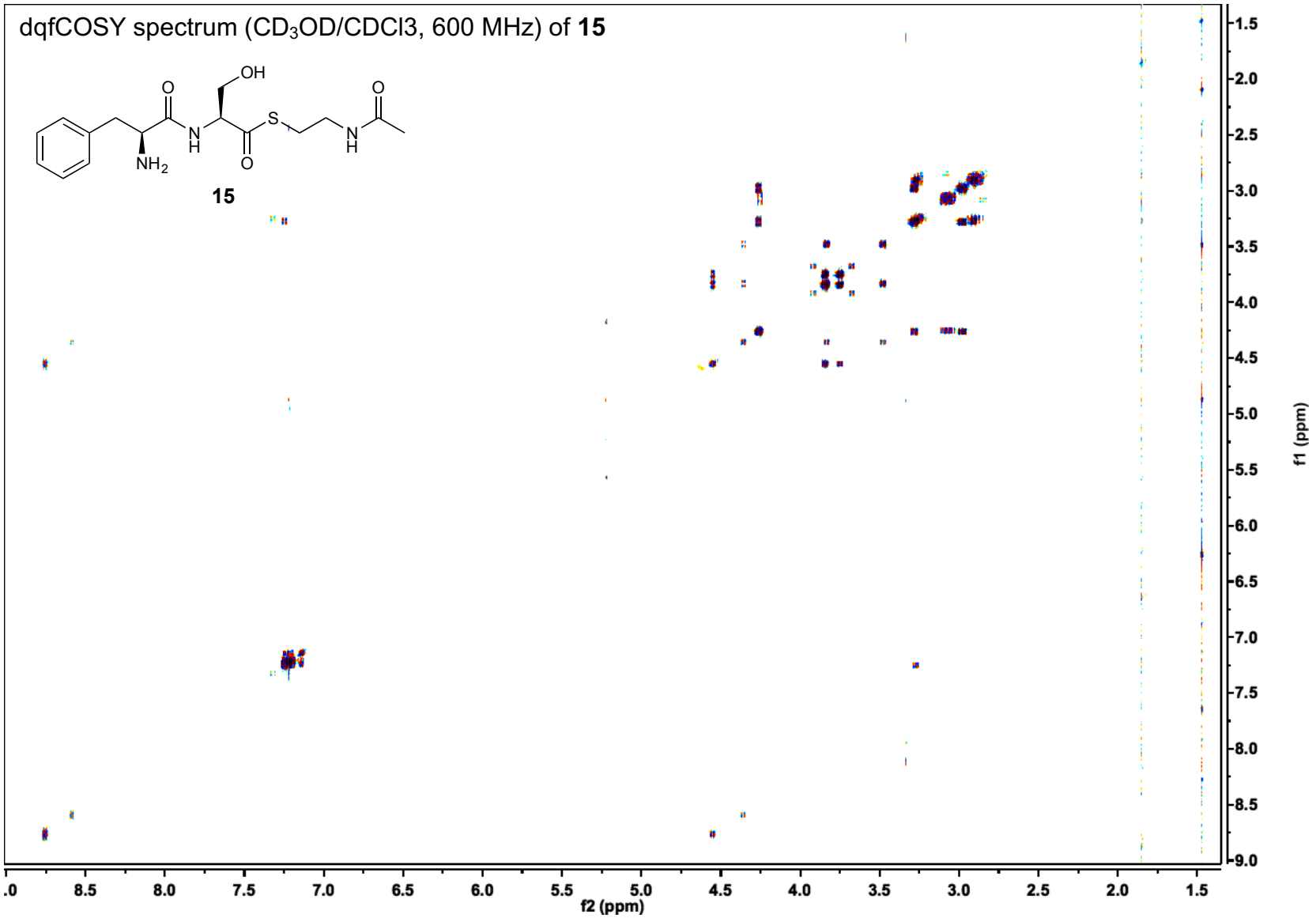
Chemical shifts were referenced to δ(C<u>H</u>D$_2$OD) = 3.31 ppm and δ(^{13}<u>C</u>HD$_2$OD$_3$) = 49.00.^{13}C chemical shifts were determined via HMBC, HSQC and direct observation ^{13}C spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



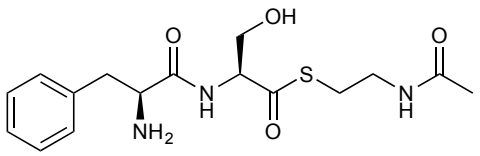
15

No.	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1	129.28	1-H	7.24 ($J_{1,2}$ = 7.5)	3
2	127.68	2-H_2	7.21 ($J_{2,1}$ = 7.5) ($J_{2,3}$ = 7.4)	2,4
3	128.92	3-H_2	7.25 ($J_{3,2}$ = 7.4) ($J_{3,5}$ = 1.0)	1,3,5
4	133.65			
5	36.87	5-H_a	2.98 ($J_{5a,5b}$ = 14.0) ($J_{5a,6}$ = 8.2) ($J_{5a,3}$ = 1.0)	3,4,6,7
		5-H_b	3.28 ($J_{5b,5a}$ = 14.0) ($J_{5b,6}$ = 5.5) ($J_{5a,3}$ = 1.0)	3,4,6,7
6	54.39	6-H	4.26 ($J_{6,5a}$ = 8.2) ($J_{6,5a}$ = 5.5)	4,5,7
7	168.86			
8		8-NH		
9	61.54	9-H	4.55 ($J_{9,10a}$ = 4.0) ($J_{9,10b}$ = 4.5)	7,10,11
10	61.70	10-H_a	3.75 ($J_{10a,10b}$ = 11.8) ($J_{10a,9}$ = 4.0)	9,11
		10-H_b	3.84 ($J_{10b,10a}$ = 11.8) ($J_{10b,9}$ = 4.5)	9,11
11	198.19			
12				
13	28.44	13-H_2	2.90 ($J_{13,14}$ = 12.0)	11,14
14	38.48	15-H_2	3.27 ($J_{14,13}$ = 12.0)	13,16
15		15-NH		
16	172.01			
17	22.31	17-H_3	1.86	16


1H NMR spectrum (CD3OD/CDCl3, 600 MHz) of 15
OH
O
N
H
S
O
N
H
O
NH2
15
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)

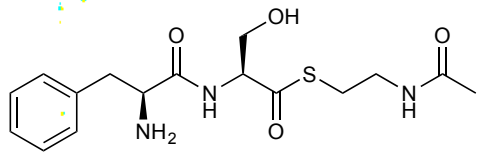

dqfCOSY spectrum (CD3OD/CDCl3, 600 MHz) of 15
OH
O
O
S
N
H
N
H
NH2
O
15
f1 (ppm)
f2 (ppm)

HSQC spectrum (CD_3OD/CDCl3, 600 MHz) of **15**



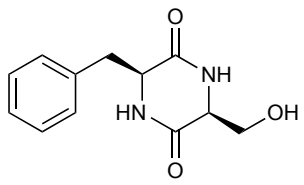
15

HMBC spectrum (CD_3OD/CDCl3, 600 MHz) of **15**



15

^{1}H NMR spectrum (Acetonitrile-d_3, 600 MHz) of **2** with Ethyl acetate solvent impurity



2

REFERENCES

1. Forseth, R. R. *et al.* Identification of cryptic products of the gliotoxin gene cluster using NMR-based comparative metabolomics and a model for gliotoxin biosynthesis. *J Am Chem Soc* **133,** 9678–9681, (2011).

2. Gao, X. *et al.* Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. *Nat Chem Biol* 1–8, (2012).

3. Forseth, R. R. *et al.* Homologous NRPS-like Gene Clusters Mediate Redundant Small-Molecule Biosynthesis in Aspergillus flavus. *Angew Chem Int Ed* **52,** 1590–1594, (2012).

4. Baccile, J. A. *et al.* Plant-like biosynthesis of isoquinoline alkaloids in Aspergillus fumigatus. *Nat Chem Biol* **12**, 419-424, (2016)**.**

5. Yu, J.-H. *et al.* Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi. *Fungal Genet Biol* **41,** 973–981, (2004).

6. Mamiatis, T., Fritsch, E. F., Sambrook, J. & Engel, J. Molecular cloning–A laboratory manual. New York: Cold Spring Harbor Laboratory. 1982, 545, (1985).

7. Bok, J. W. *et al.* Chromatin-level regulation of biosynthetic gene clusters. *Nat Chem Biol* **5,** 462–464, (2009).

8. Due-Hansen, M. E. *et al.* A protocol for amide bond formation with electron deficient amines and sterically hindered substrates. *Org Biomol Chem* **14,** 430–433, (2016).

APPENDIX C

PLANT-LIKE ISOQUINOLINE ALKALOID BIOSYNTHESIS IN *ASPERGILLUS FUMIGATUS*

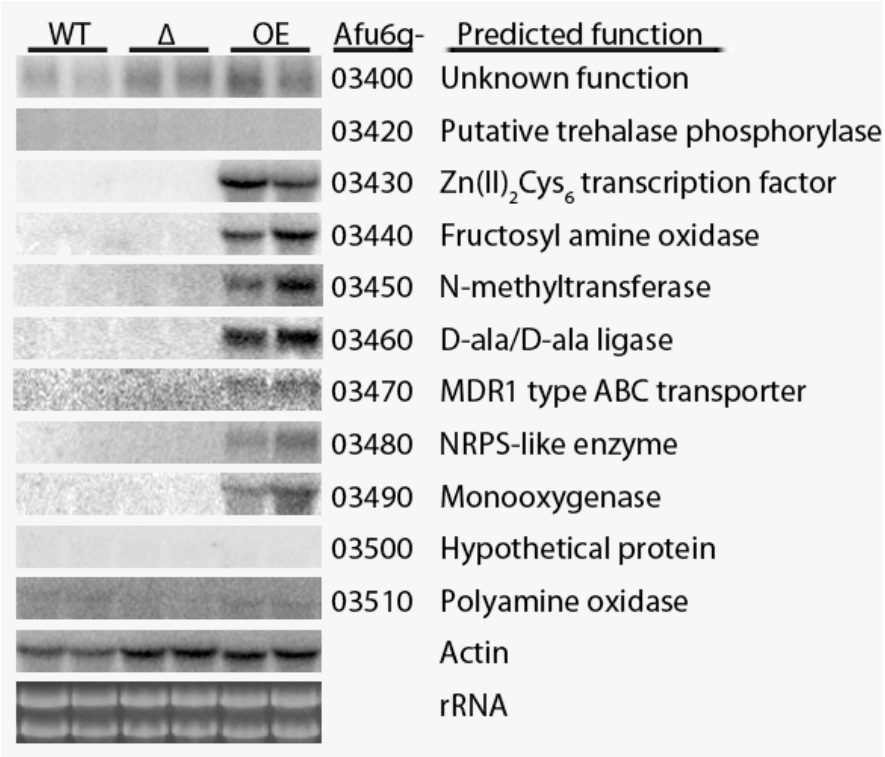


Figure C1. Northern blot analysis of WT (Af293) left, Δ*fsqA* middle, and OE::*fsqA,* right. Overexpression of *fsqA* causes specific up-regulation of AFUA6g_03430 – AFUA6g_03490, defining the boundaries of the *fsq* cluster.

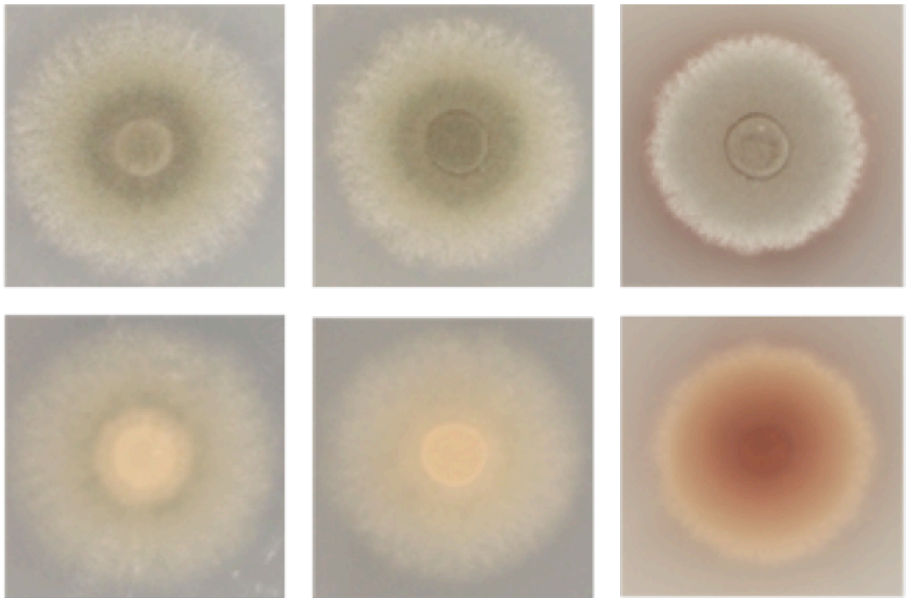


Figure C2. Phenotype of WT (Af293) left, Δ*fsqA* middle, and OE::*fsqA* right grown on GMM at 37ºC for 72 hours. OE::*fsqA* decreases radial growth and shows characteristic brown pigmentation diffusing into the media.

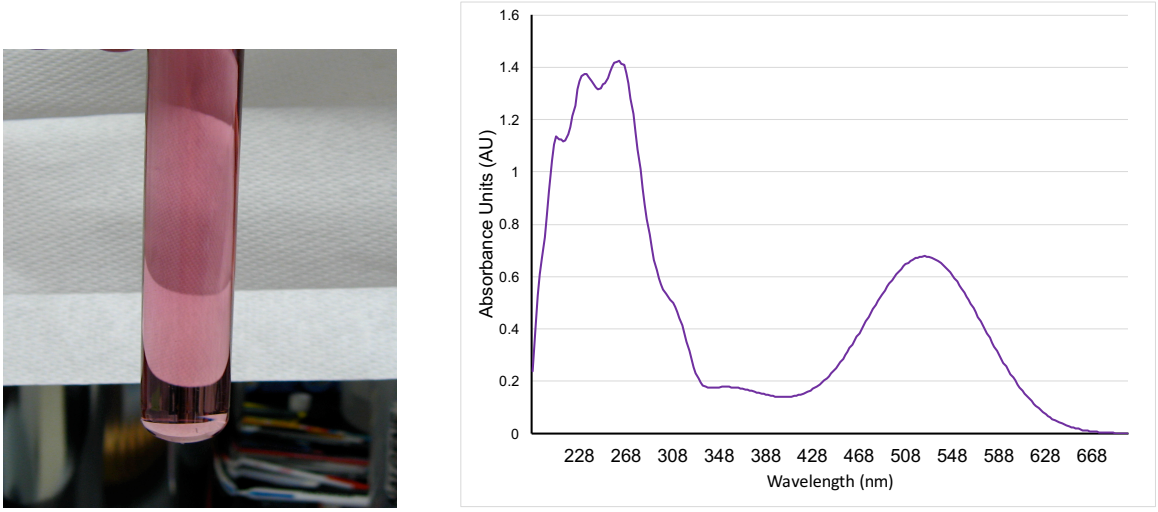


Figure C3. **Color of fumisoquin C.** Photograph of a test tube (18 x 150 mm) from large-scale reverse-phase chromatography containing ~0.02 mg/mL of fumisoquin C, **3**, (left) and UV-Vis spectrum from HPLC-UV-MS analysis (right).

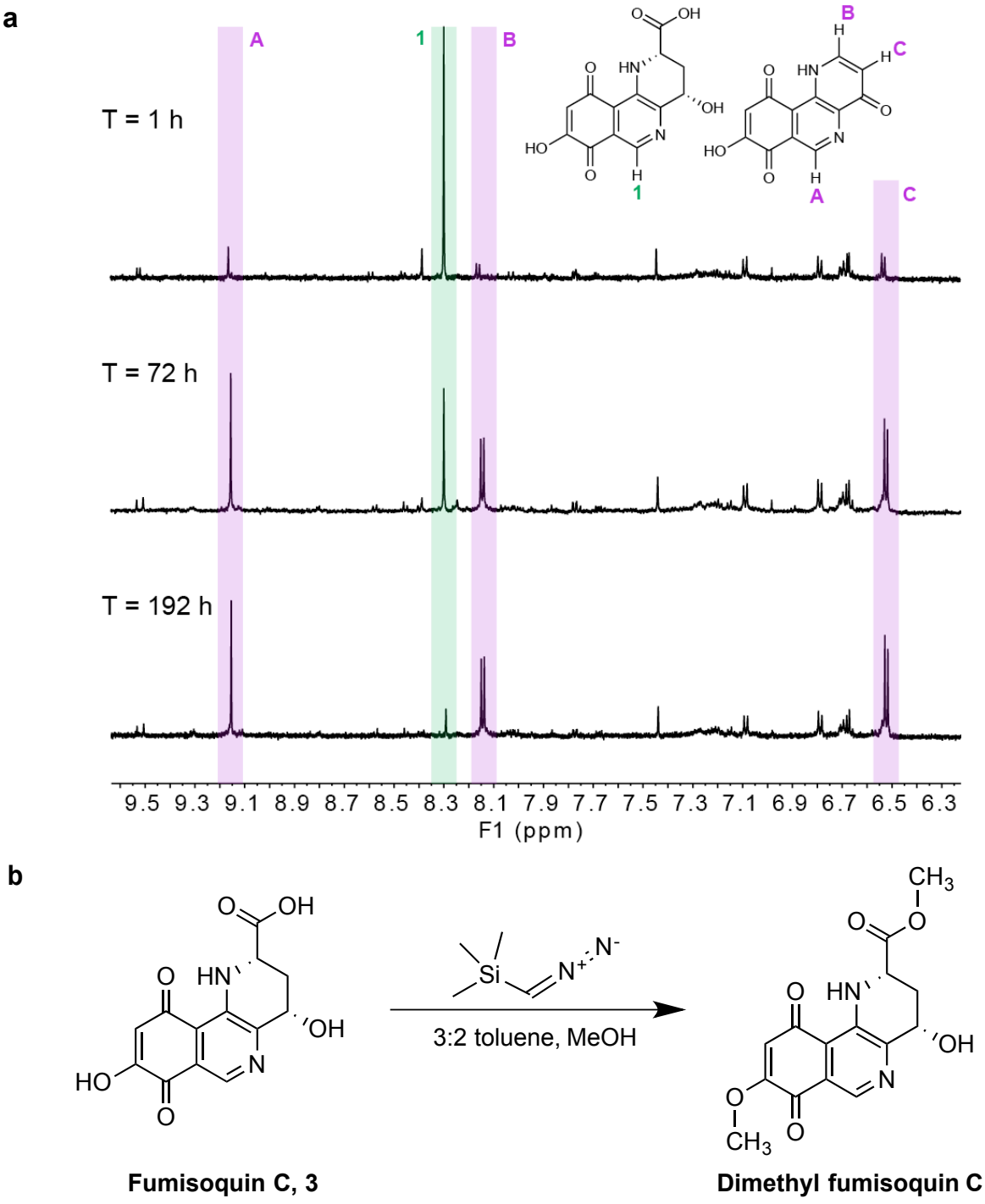


Figure C4. **Fumisoquin C decomposition and derivatization.** (**a**) ^{1}H NMR spectra in CD_3OD at T = 2 h (top), T = 72 h (middle), and T = 192 h (bottom) after chromatographic purification, showing conversion of fumisoquin C, **3**, into **4**. (**b**) Conversion of fumisoquin C (**3**) into dimethyl fumisoquin C**.** See Methods below for experimental procedure.

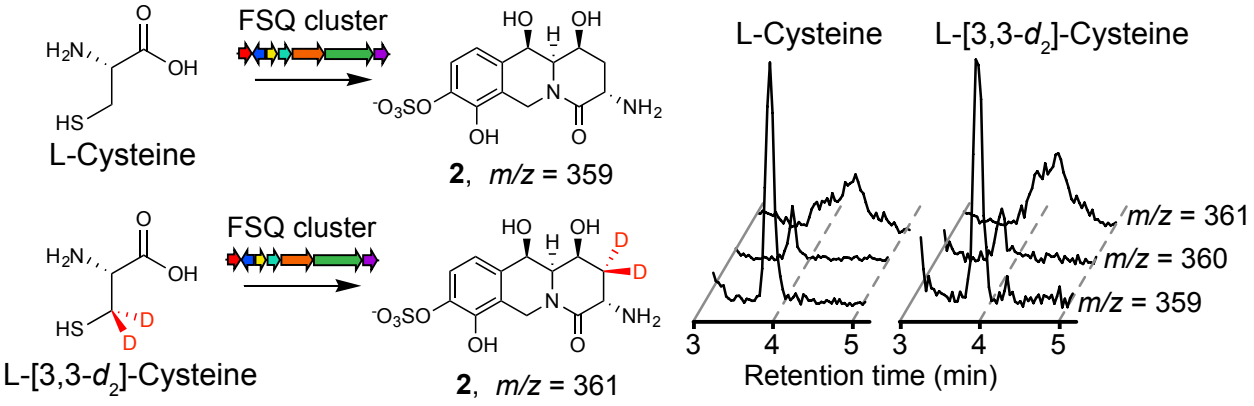


Figure C5. **L-cysteine is not incorporated into the fumisoquins.** Ion chromatograms from HPLC-MS analysis of extracts from OE::*fsqA* fed with L-cysteine or L-[3,3-d_2]-cysteine. Indicated *m/z* values correspond to isotopomers of **2** with or without heavy atom incorporation.

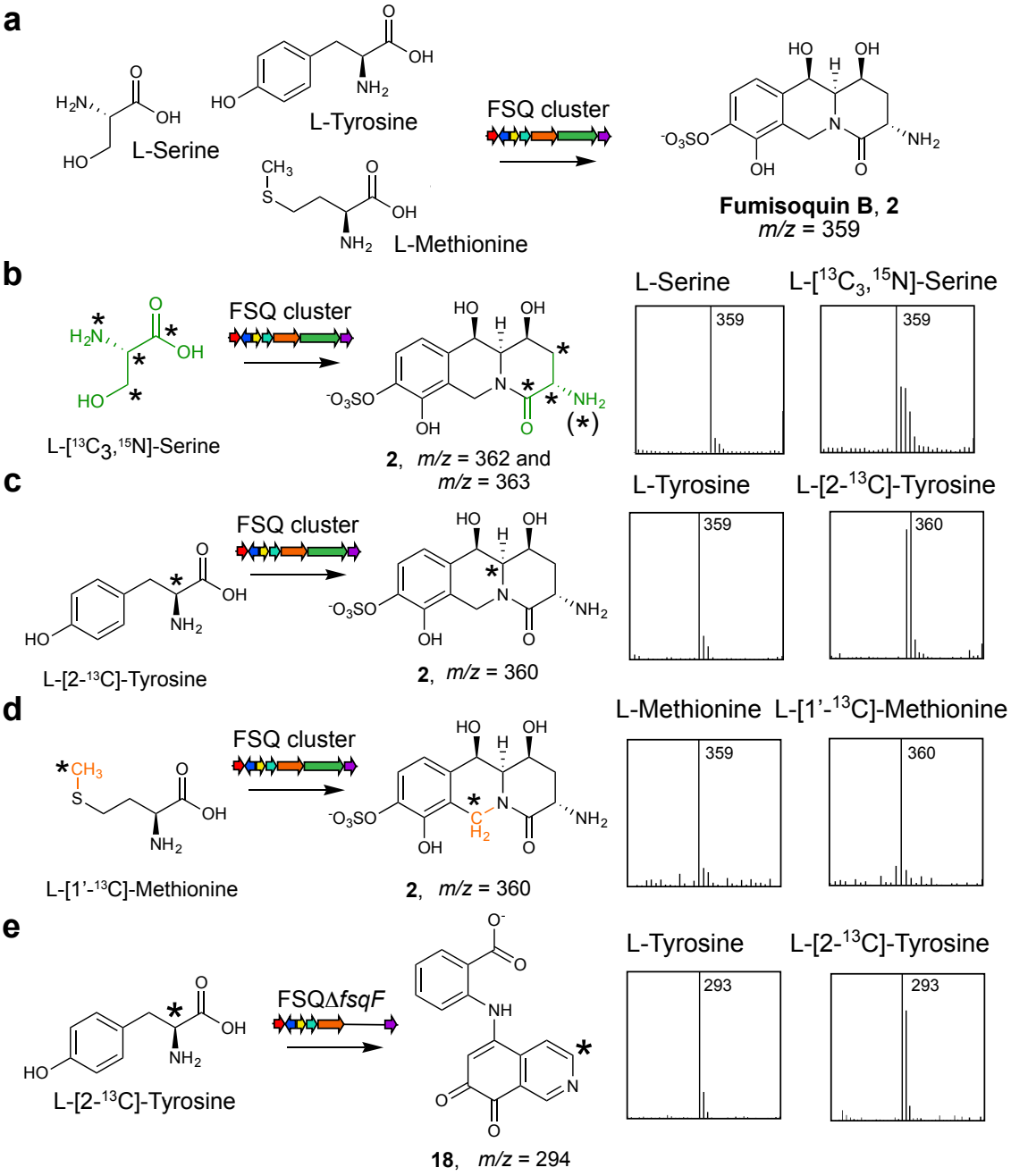


Figure C6. **Mass spectra of FSQ cluster metabolites with and without heavy atom incorporation**. (**a**) L-serine, L-tyrosine, and L-methionine are incorporated into the fumisoquins. (**b**) Mass spectrum of **2** with L-serine incorporated (left) and mass spectrum of **2** with L-[$^{13}C_3$, ^{15}N]-serine incorporated (right). Increased M+1 and M+2 peaks likely result from amination/deamination and use of abundant labeled serine for methionine production. (**c**) Mass spectrum of **2** with L-tyrosine incorporated (left) and mass spectrum of **2** with L-[2-^{13}C]-tyrosine incorporated (right). (**d**) Mass spectrum of **2** with L-methionine incorporated (left) and MS-spectrum of **2** with L-[1'-^{13}C]-methionine incorporated (right). (**e**) Mass spectrum of **18** with L-tyrosine incorporated (left) and mass spectrum of **18** with L-[2-^{13}C]-tyrosine incorporated (right).

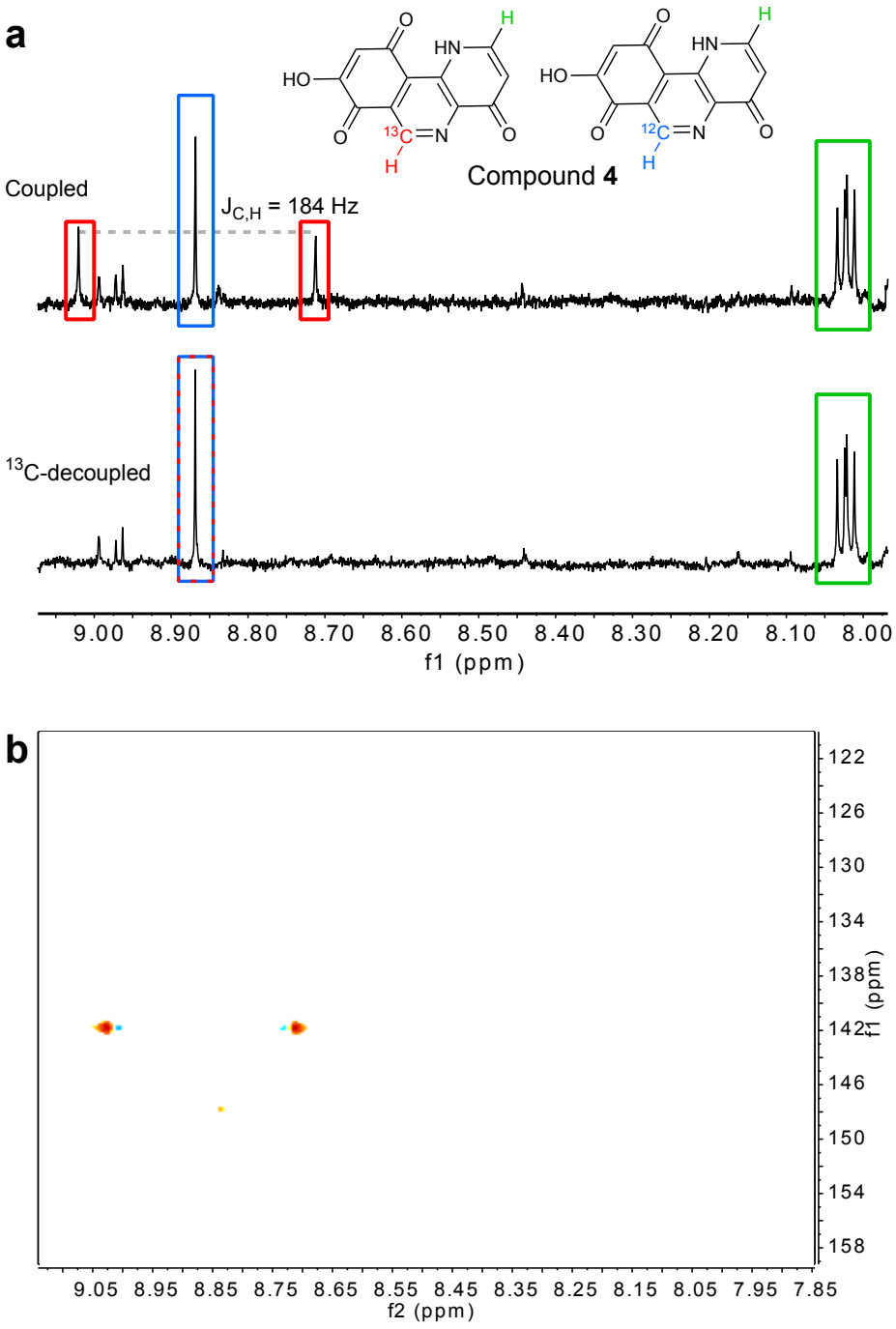


Figure C7. **NMR spectra of ^{13}C-enriched 4** (a) ^{1}H-^{13}C coupled ^{1}H NMR spectrum of a sample of **4** obtained from fumisoquin C (**3**) isolated from a fungal culture grown with L-[1'-^{13}C]-methionine (top) and ^{1}H-^{13}C *decoupled* ^{1}H NMR spectrum of the same sample of **4** (bottom), showing selective incorporation of the methionine methyl group. In the decoupled spectrum, the intensity of the signal of the proton attached to the labeled carbon does not increase proportionally due to partial signal loss and line shape changes during decoupling. (**b**) ^{1}H-^{13}C coupled HSQC spectrum of this sample of **4**. Spectra were acquired using the 600 MHz Varian INOVA spectrometer, using DMSO-d_6 as solvent.

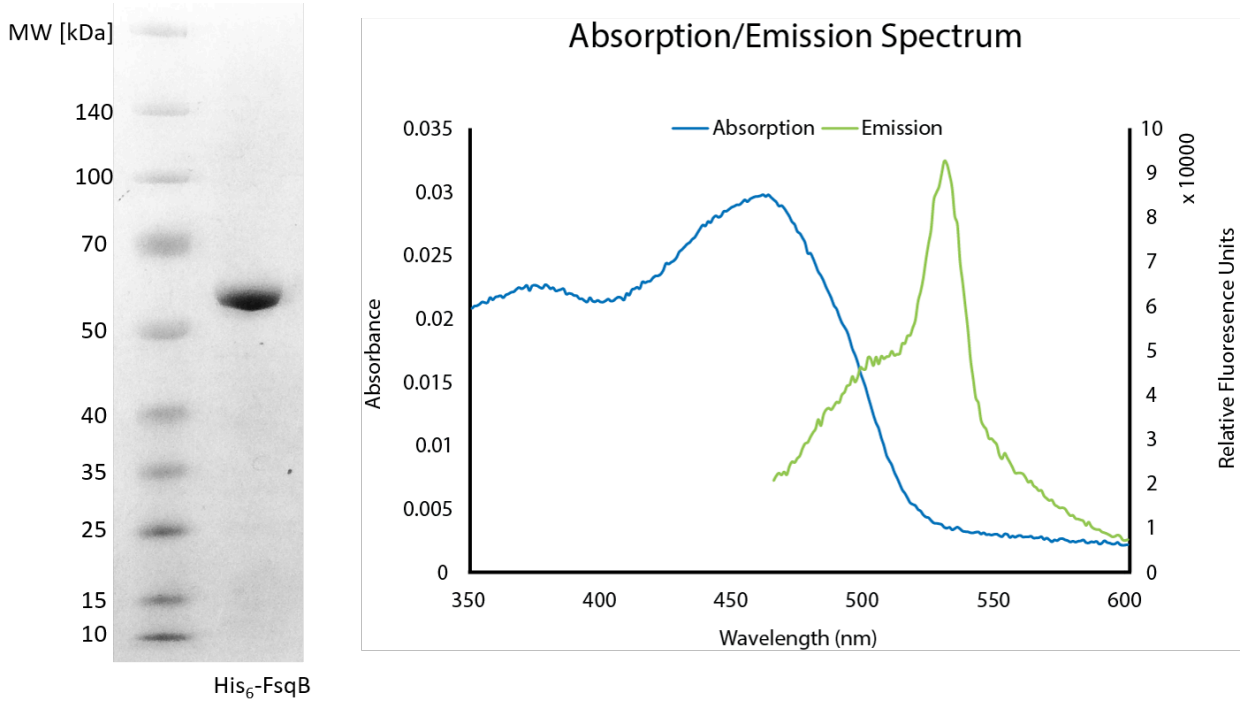


Figure C8. **Purification and spectra of FsqB.** SDS-polyacrylamide gel of the polyhistidine-tagged FsqB (left), and UV-Vis absorption and fluorescence spectra (right) of FsqB. Spectra were acquired in 100 mM potassium phosphate buffer at pH 7.0, and fluorescence spectrum was acquired using an excitation wavelength of 450 nm.

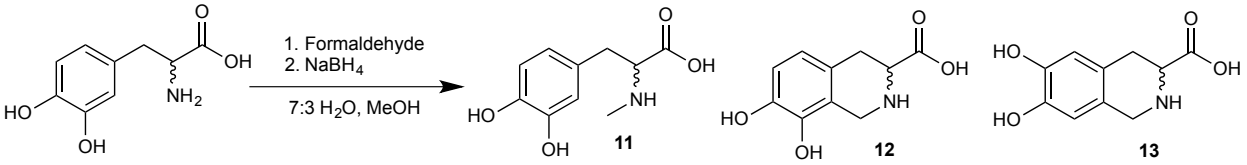


Figure C9. Reaction of DL-DOPA with formaldehyde followed by addition of sodium borohydride produces a 7:3 mixture of cyclized products **12** and .**13**, in addition to variable amounts of uncyclized **11**, as determined by HPLC-MS**.**

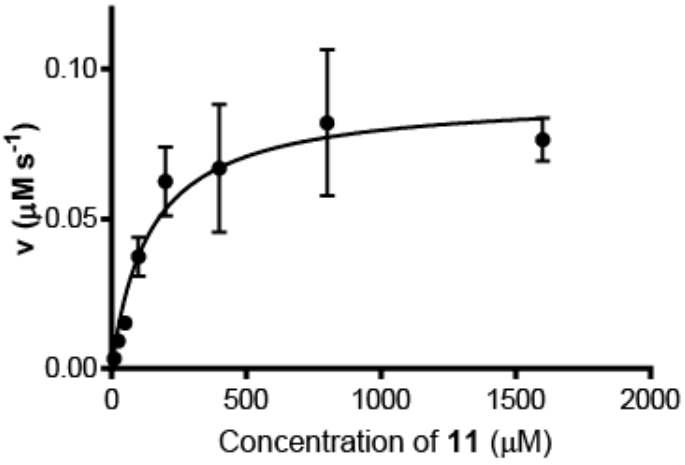


Figure C10. **Enzymatic activity of FsqB.** Steady-state kinetics evaluated for FsqB on model substrate **11**. The observed apparent steady-state kinetic parameters of FsqB operating on **11** were K_M: 142.5 ± 42.9 μM and k_{cat}: 0.9 ± 0.1 s^{-1} at 25 °C. Each initial concentration of **11** was sampled twice for kinetic analysis, and displayed as the mean ± s.d.

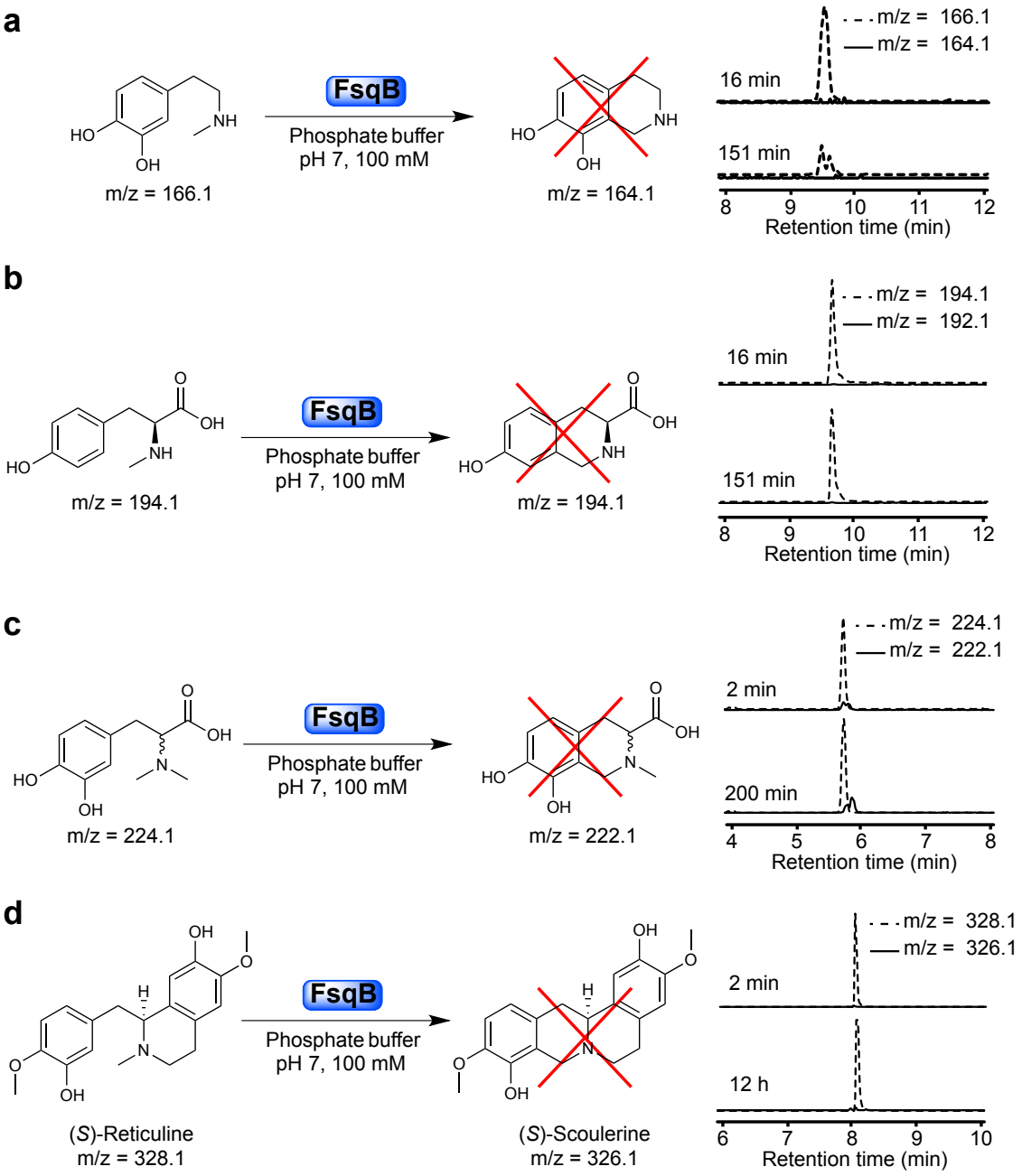


Figure C11. Substrate specificity of FsqB. (**a**) Ion chromatograms for *N*-methyl dopamine (*m/z* = 166.1) and cyclic product (*m/z* = 164.1), showing no product formation after 151 min. (**b**) FsqB does not catalyze the cyclization of *N*-methyl-L-tyrosine. Ion chromatograms for *N*-methyl-L-tyrosine (*m/z* = 194.1) and cyclic product (*m/z* = 192.1) obtained after 151 min show no product formation. (**c**) FsqB does not catalyze the cyclization of *N,N*-dimethyl DOPA. Ion chromatograms for *N,N*-dimethyl DOPA (*m/z* = 224.1) and putative cyclic product (*m/z* = 222.1) obtained after 200 min show no product formation. Validating FsqB preference toward secondary β-*N*-methylamine substrates. (**d**) FsqB does not catalyze the cyclization of (*S*)-reticuline. Ion chromatograms for (*S*)-reticuline (*m/z* = 328.1) and cyclic product (*m/z* = 326.1) obtained after 12 h show no product formation.

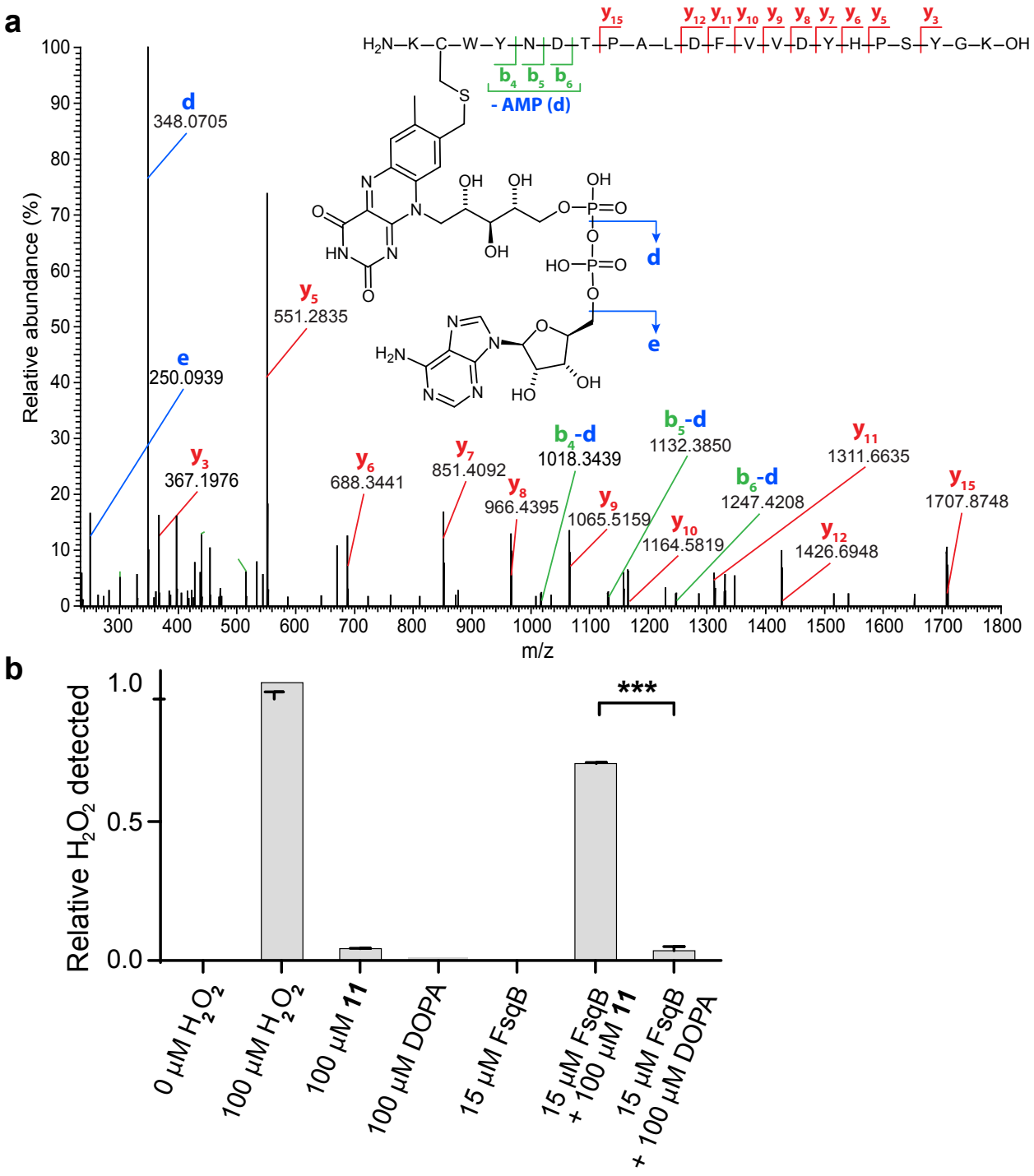


Figure C12. **FsqB has covalently bound FAD cofactor and releases H_2O_2 in the presence of substrate** (**a**) Collision-induced dissociation mass spectrum of tryptic FsqB peptide, revealing covalently bound FAD. Amino acid sequence and location of FAD attachment were inferred from indicated *b*- (green) and *y*-type (red) ion series and diagnostic FAD fragmentation (blue). (**b**) Amplex Red H_2O_2 assay of 15 μM FsqB with or without **11**, or 3,4-dihydroxy-DL-phenylalanine (DOPA) (100 mM phoshate buffer, pH 7, 1.5 h). Values were normalized to 0 and 100 μM H_2O_2 and presented as mean ± s.d. (n = 3).

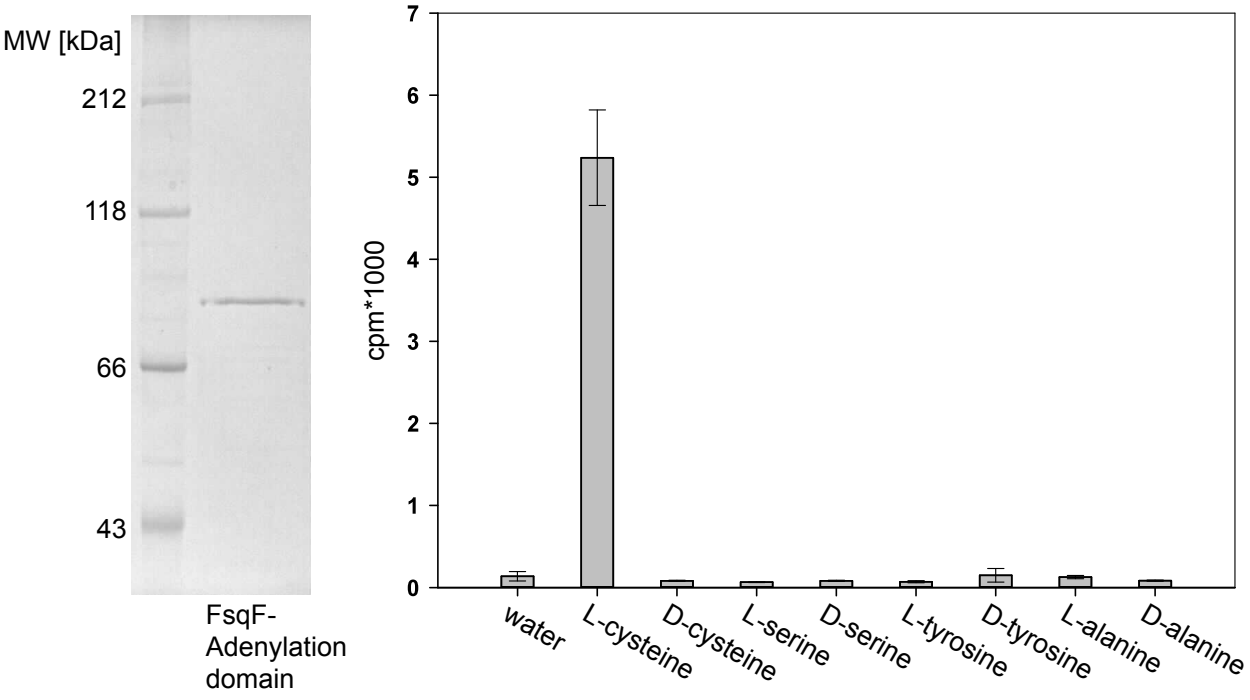


Figure C13. **Properties of recombinant FsqF adenylation domain.** SDS-polyacrylamide gel showing the purity of the FsqF adenylation domain (left), and ATP-[^{32}P]pyrophosphate exchange assay results (right) for amino acids selected based on results from the stable-isotope labeling experiments (see **Figure 3.4a-c** and **Figure C5** and Methods below). The exchange assay shows none or minimal (in the case of L-cysteine) activation for all tested amino acids. Collectively, the assay results suggest that no standard amino acid is the true substrate, but rather a derivative of an amino acid, such as dehydroalanine, as proposed in the biosynthetic model shown in main text **Figure 3.3**.

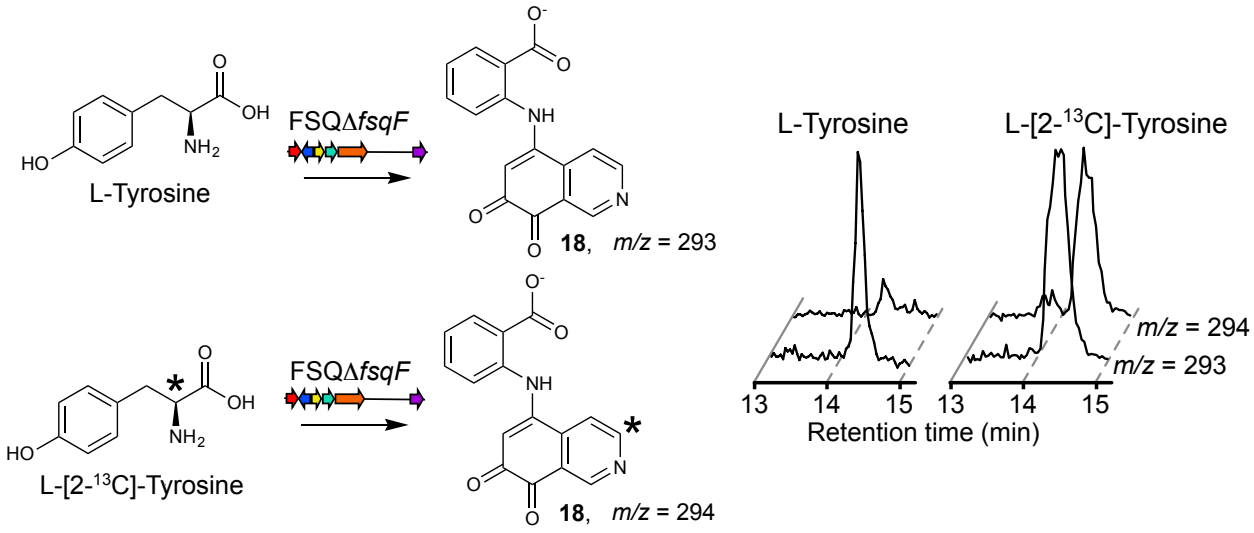


Figure C14. **L-tyrosine is incorporated into shunt metabolite 18.** Ion chromatograms extracted from HPLC-MS analysis of extracts from OE::*fsqA*-Δ*fsqF* fed with L-tyrosine or L-[2-^{13}C]-tyrosine. The indicated *m/z* values correspond to **18** with or without heavy atom incorporation. The data support the hypothesis that FsqD is responsible for incorporation of L-tyrosine and not FsqF, as shown in main text **Figure 3.3.**

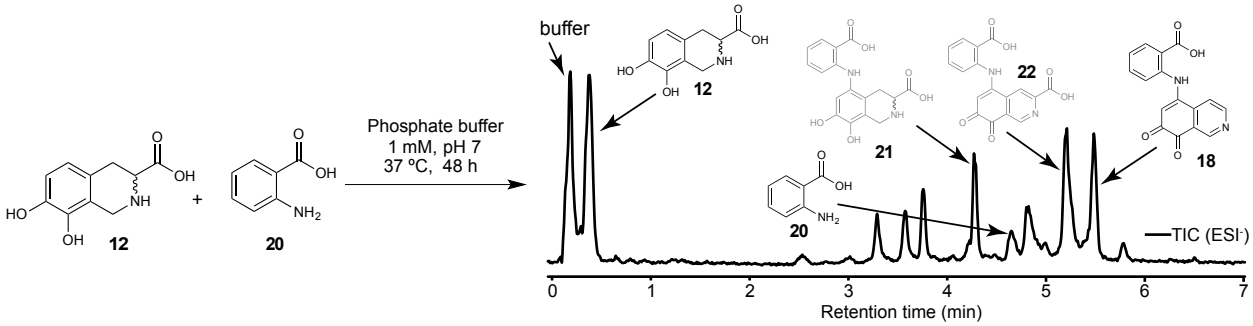


Figure C15. Shunt metabolite 18 is produced non-enzymatically from 12 and anthranilic acid, 20. High-resolution UHPLC-MS total ion chromatogram for a reaction mixture of **12** and **20** reveals formation of **18**, along with two intermediates, **21** and **22** (proposed structures shown in gray). See **Table C1** for HRMS data.

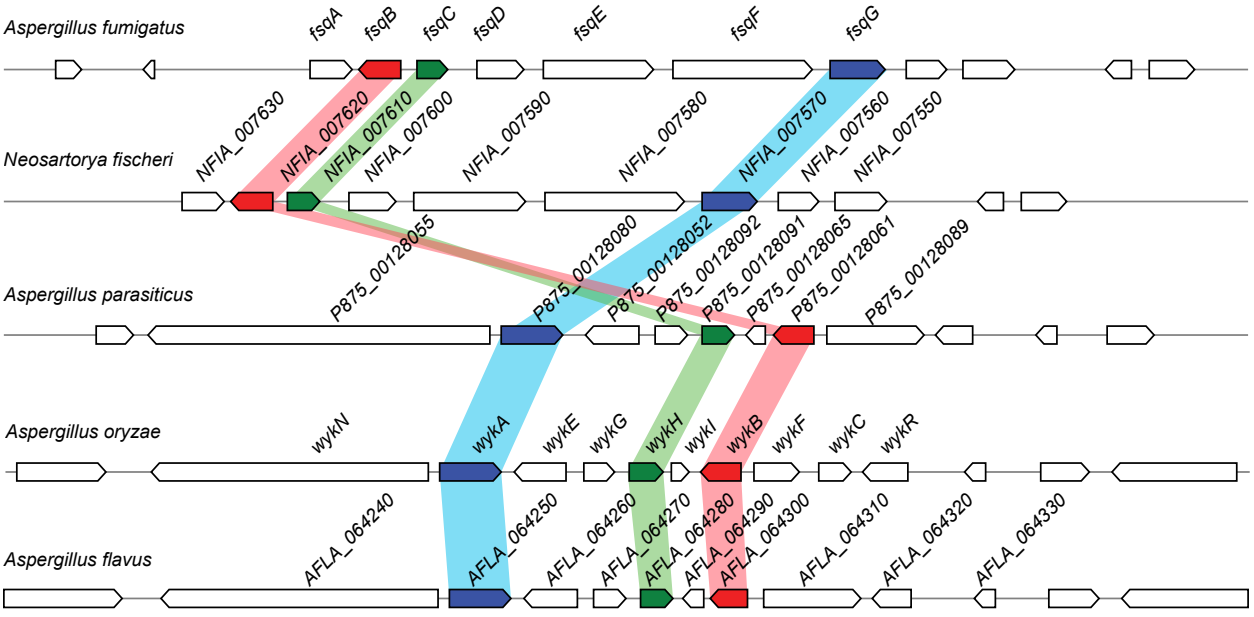


Figure C16. Synteny analysis of *fsq* cluster in indicated *Aspergillus* and *Neosartorya* species. The analysis shows conservation of *fsqB* (red), *fsqC* (green), and *fsqG* (blue), which encode the enzymes responsible for the incorporation of the isoquinoline ring in **1-5**, and **18**. For % identity of these genes in each species shown, see **Table C2**.

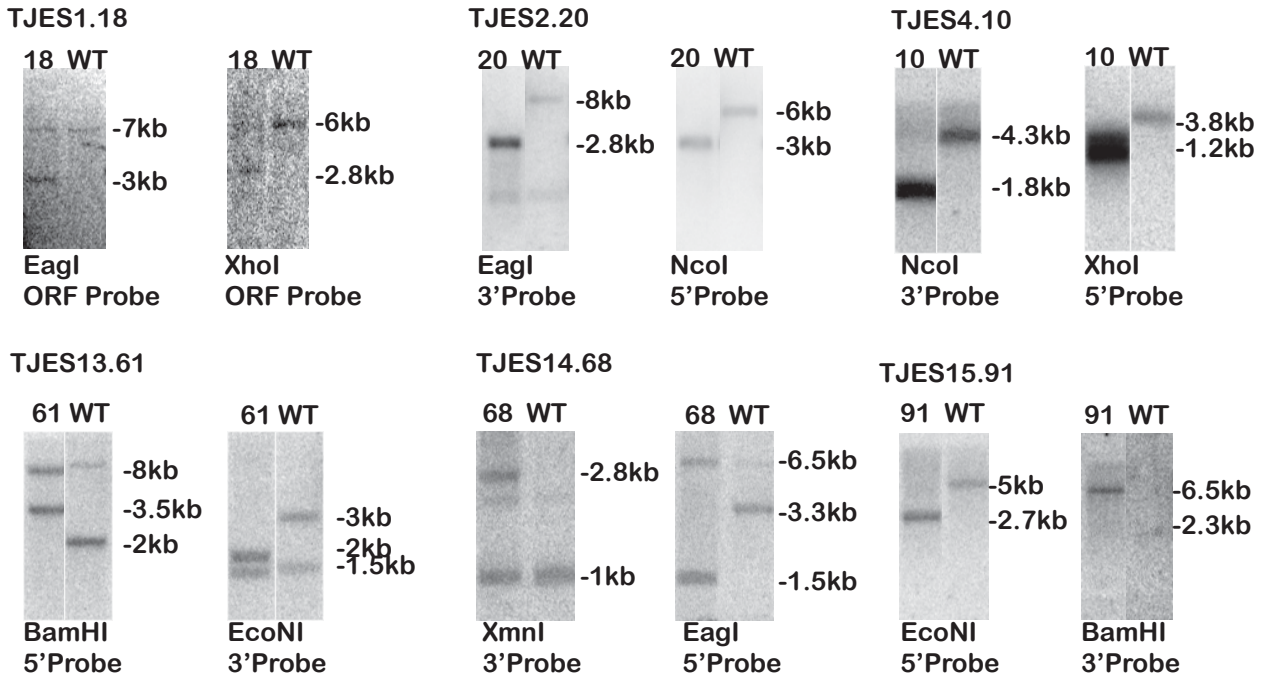


Figure C4. Southern analysis confirmation of all mutants used in this study. Mutant (number) lanes are on the left of each image and the parental strain ("WT") is shown on the right. Expected band sizes correlating with those seen in the images are marked accordingly. Under each image is the enzymes used in the restriction digest as well as the nucleic acid probe used.

Strains, media, and growth conditions. The fungal strains used in this study are listed in **Table C4**. All strains were grown at 37 °C on glucose minimal medium (GMM)[1] and, when appropriate, were supplemented with 0.56 g/L uracil, 1.26 g/L uridine, 1.0 g/L arginine and maintained as glycerol stocks at -80 ºC. *Escherichia coli* strain DH5α was propagated in LB medium with appropriate antibiotics for plasmid DNA.

Gene cloning, plasmid construction, and genetic manipulations. The plasmids utilized in this study are listed in **Table C4**. The oligonucleotide sequences for PCR primers are given in **Table C5**. PCR amplification was carried out on a C1000^{TM} Thermal Cycler (Bio-Rad). For creation of the *fsqA* overexpression (OE) strain (TJES1.18) the OE cassette was constructed by amplifying the *fsqA* open reading frame (ORF) from *Af293* genomic DNA using primers Afu6g03430_NcoI_For and Afu6g03430_NotI_Rev, which introduced a 5' NcoI restriction site and a 3' NotI restriction site. This PCR product was purified with a QIAquick gel extraction kit (Qiagen), quantified, and digested with the appropriate restriction enzymes before being cloned into pJMP8.12[2], resulting in plasmid pJES 1.2. Subsequently, *A. fumigatus argB* was transferred from pJMP4.1[3] by digesting with EcoRI and introduced into pJES1.2, resulting in plasmid pJES2.7 (*A. fumigatus argB::gpdA(p)::fsqA*). All PCR steps were carried out with Pfu Ultra II DNA Polymerases (Agilent) and all digestion reactions were carried out with NEB enzymes (New England BioLabs). Correctness of the inserted DNA was then confirmed by sequencing.

Construction of the *fsq* gene knock-out cassettes were generated using standard double-joint PCR procedures[4]. Briefly, *A. parasiticus pyrG* (*A.ppyrG)* was amplified from pJMP9.1. Then, an approximately 1 kb fragment upstream and downstream of each *fsq* gene was amplified from genomic DNA of *A. fumigatus* Af293 using designated primers, respectively. These three amplified PCR products were purified with a QIAquick gel

extraction kit, quantified, and fused using standard double-joint PCR procedures. The final PCR product was amplified with the bottom primer pairs, *gene*_5'F_flank and _3'R_flank, confirmed with endonuclease digestion, and purified for fungal transformation. The first two rounds of PCR were done with Pfu Ultra II DNA Polymerases (Agilent), and the final PCR step used Expand long template PCR system (Roche) according to the manufacturer's instructions. Af293.6 (double auxotroph *A. fumigatus, pyrG*$^{-}$*, argB*$^{-}$) was used to make the *OE*::*fsqA, pyrG*$^{-}$ auxotroph, TJES1.18. This strain was used as the recipient host strain for subsequent deletions of *fsq* genes as well as the ectopic complementation of *pyrG1* with plasmid pJMP9.1 (containing *A.ppyrG*) resulting in prototrophic strain TJES3.1 (OE::*fsqA, A.ppyrG*). Similarly Δ*fsqA* auxotrophic strains (TJES2.20) were made by transforming Af293.6 with deletion cassettes and complemented with ectopically integrated pJMP4 to make Δ*fsqA* prototroph TJES8.2.

For the creation of deletion mutants in OE*::fsqA* TJES1.18 background, a deletion cassette of each cluster gene (*fsqB* to *fsqG*) was constructed by using double-joint PCR with *A. parasiticus pyrG* (*A.ppyrG*) for replacement of target genes. The *A.ppyrG* gene was amplified from the plasmid pJMP9.1 as template. Five µg of the double-joint PCR cassette were used to delete each *fsq* cluster gene by using TJES1.18 (*A.fargB::gpdA(p):: fsqA, pyrG1*) as the recipient host. After transformation, transformants were grown on minimal media plates without supplements for screening. All strains were verified by PCR and Southern blot analysis (**Figure C17**). Multiple confirmed strains for each mutant were stored at -80 °C with 33 % glycerol for future use.

Heterologous protein production and biochemical assays. The expression vectors for the *fsqF* (NRPS-like) adenylation domain (pJES13.2) were constructed by amplifying

bases 1-1968 of the coding DNA sequence (CDS) of the gene Afu6g03480 from Af293 using primers fqsF_Adomain_xp_5'F and fqsF_Adomain_xp_3'R introducing 5' NotI and 3' XhoI restriction sites, respectively. These products were gel purified and digested with the appropriate enzymes before being cloned into pET30-a(+) (EMD Biosciences) to create an expression plasmid with a 6xHis tagged *fsqF* adenylation domain. The plasmid was introduced into chemically competent *E.coli* DH5α, selected on LB agar with Kanamycin and confirmed by sequencing. Subsequently, they were extracted from DH5α cells and used to transform expression host *E.coli* BL21. Protein purification was carried out as described[5]. All reactions were carried out in triplicate. Reaction parameters for the ATP-[^{32}P]pyrophosphate exchange assay were: total assay volume of 100 µL at 25°C in 100 mM phosphate buffer, 5 mM $MgCl_2$, 125 nM EDTA, 5 mM ATP, 100 nM purified FsqF adenylation domain, 0.1 µM [^{32}P]pyrophosphate (50 Ci/mmol), and 1 mM amino acid substrate. The reaction proceeded for 30 min before it was stopped (1 % (w/v) activated charcoal, 4.5 % (w/v) tetrasodium pyrophosphate, 3.5 % (v/v) perchloric acid) and further processed as described[6]. Pyrophosphate exchange was quantified on a scintillation counter (PerkinElmer TriCarb 2910TR).

The expression vector for *fsqB* (fructosyl amino acid oxidase) was constructed using the CDS of the gene Afu6g03440 from OE::*fsqA* using primers pet28_fsqB_3'_fwd and fsqB_5'_pet28_rev. These products were gel purified and cloned into pET28-b(+) using ligase-free PCR cloning to create an expression plasmid with a 6xHis tagged FsqB. The plasmid was introduced into chemically competent *E. coli* DH5α, selected on LB agar with Kanamycin and confirmed by sequencing. Subsequently, they were extracted from DH5α cells and used to transform expression host *E. coli* BL21(DE3) (New England BioLabs). BL21(DE3) pET28-b(+) *fsqB* containing *E. coli* were grown in LB containing 50 µg/mL Kanamycin at 37 °C to an OD of ~0.6, cooled to 18 °C and expression was

induced with 250 μM IPTG for 15 hrs. 4 L of LB containing *E. coli* were harvested at 6000 G for 20 min yielding 10.6 g of wet weight pellet and stored at -80 °C. Protein purification was carried out at 4 °C by resuspending the pellet in lysis buffer containing 100 mM phosphate buffer pH 7.8, 10 mM imidazole, 10 % (v/v) glycerol, 1 mM PMSF, 0.1 mg/mL lysozyme, 0.1 mg/mL Benzonase (EMD Millipore) and sonicated. All following steps were performed in low light to avoid detrimental flavin mediated reactions. Lysed cells were centrifuged at 20000 G for 20 min, and loaded onto TALON affinity resin (Clontech) pre-equilibrated with lysis buffer and incubated for 30 min with stirring. Resin was loaded onto a column and washed with 2 column volumes of lysis buffer without lysosome or Benzonase. FsqB was eluted with 100 mM phosphate buffer pH 7.8, 150 mM imidazole, 10 % (v/v) glycerol with 8 column volumes, concentrated and buffer exhanged with Amicon Ultra-15 centrifugal filter 30 kDa cutoff with 100 mM phosphate buffer pH 7.8, 25 % (v/v) glycerol and stored at -80°C. Activity assays were performed in total assay volumes of 100 μL at 37 °C in 100 mM phosphate buffer pH 7.0 in the dark. Reaction products were identified on a Thermo Scientific-Dionex Ultimate3000 UHPLC system equipped with a diode array detector and connected to a Thermo Scientific Q Exactive Orbitrap operated in electrospray negative (ESI-) ionization mode. UV-Vis absorbance spectra were obtained on a Cary 5000 UV-Vis-NIR spectrophotometer, and fluorescence emission was obtained on a HORIBA Nanolog Spectrofluorometer.

FsqB Steady-state kinetic analysis and imine capture. Steady-state kinetics were carried out at 25 °C in 100 mM phosphate buffer pH 7.0 with 100 nM FsqB at final volumes of 500 or 1000 μL at various initial concentrations of *N*-methyl-3,4-dihydroxy-DL-phenylalanine (**11**). Imine capture was carried out at 25 °C in 100 mM phosphate buffer pH 7.0 with 1.5 μM FsqB, 400 μM *N*-methyl-L-tyrosine (**14**), and 1 mM dimedone

(**15**) (Sigma Aldrich) with <0.5% (v/v) methanol for dimedone solubility. Both kinetics and imine capture were monitored with low-resolution HPLC-MS performed on an Agilent 1100 series HPLC system equipped with a diode array detector and connected to a Quattro II mass spectrometer (Micromass/Waters) operated in electrospray negative ionization (ESI^-) mode for kinetic analysis and positive electrospray ionization (ESI^+) mode for imine capture. Kinetic data analysis was performed using GraphPad Prism version 6.00 for Windows (GraphPad Software).

FsqB Amplex Red assay. 100 μM *N*-methyl-DL-DOPA or DL-DOPA, either with or without 15 μM FsqB, were incubated in 100 mM pH 7.0 potassium phosphate buffer at 25 °C in 100 μL total reaction volume in the dark in a 96-well plate with gentle shaking. After 1.5 h of incubation, 100 μL of hydrogen peroxide detection mixture (containing 50 mM potassium phosphate pH 7.0, 100 μM Amplex Red (Thermo Scientific), 0.2 units/mL HRP) was added and incubated at 25 °C in the dark for 10 minutes. Readings were performed on a BioTek Synergy 2 96-well plate reader using a 500±27 nm excitation filter and a 615±15 nm emission filter and normalized to 0 and 100 μM hydrogen peroxide standards. All conditions were performed in triplicate.

FsqB trypsinization analysis. 80 μg of recombinant FsqB was buffer exchanged into 55 μL of 50 mM ammonium bicarbonate. 3 μL of 0.33 μg/μL of sequencing grade porcine trypsin (Promega) was added to FsqB and incubated at 37 °C for 12 h. 5 μL 50 mM TCEP was added and the mixture was further incubated at 37 °C for 10 min. The solution was cooled to 25 °C and 5 uL of 100 mM iodoacetamide was added and incubated at 25 °C in the dark for 1 h. Formic acid was added to 1 % (v/v), and peptides were prepared with 100 μL Pierce C_{18} Tips (Thermo Scientific) per manufacturers protocol utilizing 100 μL elution solution. Peptides were identified via UHPLC-MS/MS using a Thermo Scientific Dionex Ultimate 3000 UHPLC system equipped with a diode

array detector and connected to a Thermo Scientific Q Exactive Orbitrap operated in positive electrospray (ESI^+) ionization mode.

Nucleic acid analysis. Preparation of plasmids, restriction enzyme digestions, gel electrophoresis, blotting, probe preparation and hybridization were carried out by standard protocols. *Aspergillus* DNA was extracted using a previously described method[7]. Sequence data were analyzed using the LASERGENE software package from DNASTAR.

Northern analysis. 50 mL of liquid GMM[1] were inoculated with 1.0×10^6 spores (asexual) mL^{-1} of all appropriate strains in this study and incubated with shaking at 250 rpm at 25 °C. After 48 h, the mycelium was collected and total RNA was extracted by using Isol-RNA Lysis Reagent according to the manufacturer's instructions (5 Prime).

In order to determine the boundaries of the *fsq* gene cluster, gene fragments of potential cluster genes used as probes were amplified individually from Af293 genomic DNA with appropriate primers. WT, OE::*fsqA* (TJES3.1), and Δ*fsqA* (TJES8.2) were used for this experiment. About 40 µg of total RNA were used for RNA blot analysis. RNA blots were hybridized with designated DNA fragments. All experiments were performed in duplicate. Detection of signals was carried out with a Storm 860 phosphoimager (Molecular Dynamics).

Fermentation and metabolome extraction for comparative metabolomics by DANS and HPLC-UV-MS. Preparation for NMR spectroscopic analysis: *A*. *fumigatus* strains were inoculated (1.0×10^6 spores mL^{-1}) into 1 L GMM[1] in a 2 L Erlenmeyer flask at 37 °C with shaking at 220 rpm. After 4 days, liquid fungal cultures including fungal tissue and media were frozen using a dry ice acetone bath, and lyophilized. The lyophilized residues were extracted with 500 mL of 10 % methanol in ethyl acetate for 3.5 h with vigorous stirring. Extracts were filtered over cotton, evaporated to dryness, and stored in

8 mL glass vials at -20 °C. Prior to NMR spectroscopic analysis, the crude extracts (~30-50 mg) were suspended in 0.15 mL of methanol-d_4. The resulting suspension was evaporated to dryness, the residue re-suspended in 0.6 mL of methanol-d_4, centrifuged to remove insoluble materials, and the supernatant was transferred into a 5 mm NMR tube.

Preparation for HPLC-MS analysis: *A*. *fumigatus* strains were inoculated (1.0×10^6 spores mL^{-1}) into 50 mL GMM[1] in a 125 mL Erlenmeyer flask at 37 °C with shaking at 220 rpm. After 4 days, liquid fungal cultures including fungal tissue and media were frozen using a dry ice acetone bath, and lyophilized. The lyophilized residues were extracted with 30 mL of MeOH for 1.5 h with vigorous stirring. Extracts were filtered over cotton, evaporated to dryness, and stored in 4 mL vials. Crude extracts were suspended in 0.5 mL of MeOH and centrifuged to remove insoluble materials, and the supernatant was subjected to HPLC-MS analysis.

Fermentation and metabolome extraction for compounds 3-5.

A. *fumigatus* strains were inoculated (1.0×10^6 spores mL^{-1}) into 1 L GMM[1] in a 2 L Erlenmeyer flask at 37 °C with shaking at 220 rpm. After 3.5 days, liquid fungal cultures including fungal tissue and media were frozen using a dry ice acetone bath, and lyophilized. The lyophilized residues were extracted with 500 mL of MeOH for 1 h with vigorous stirring. Extracts were filtered over cotton and evaporated on Celite-545 (Acros Organics) to dryness. The dry Celite was then loaded into a 25 gram solid phase cartridge (Isco) and subjected to large-scale reverse phase chromatography (see Chromatographic enrichment section below for details). Compounds **4** and **5** are also minor components of the culture media.

Analytical methods and equipment overview. (a) NMR spectroscopy: NMR spectroscopic instrumentation: a Varian INOVA 600 MHz NMR spectrometer (600 MHz

^{1}H reference frequency, 151 MHz for ^{13}C) equipped with an HCN indirect-detection probe or a Bruker Avance^{III} HD (800 MHz ^{1}H reference frequency, 201 MHz for ^{13}C) equipped with a 5 mm CPTCL ^{1}H-^{13}C/^{15}N cryo probe. Non-gradient phase-cycled dqfCOSY spectra were acquired using the following parameters: 0.6 s acquisition time, 400-600 complex increments, 8, 16 or 32 scans per increment. Non-gradient HSQC, HMQC, and HMBC spectra were acquired with these parameters: 0.25 s acquisition time, 200-500 increments, 8-64 scans per increment. ^{1}H, ^{13}C-HMBC spectra were optimized for $J_{H,C}$ = 6 Hz. HSQC spectra were acquired with or without decoupling. Susceptibility-matched NMR tubes (Shigemi) were used for sample amounts smaller than 1 mg. NMR spectra were processed and baseline corrected using MestreLabs MNOVA software packages. (b) Mass spectrometry: high-resolution UHPLC-MS was performed on a Thermo Scientific-Dionex Ultimate3000 UHPLC system equipped with a diode array detector and connected to a Thermo Scientific Q Exactive Orbitrap operated in electrospray positive (ESI^{+}) or electrospray negative (ESI^{-}) ionization mode. Low-resolution HPLC-MS was performed on an Agilent 1100 series HPLC system equipped with a diode array detector and connected to a Quattro II mass spectrometer (Micromass/Waters) operated in ESI^{+} or ESI^{-} mode. Data acquisition and processing for the LC-HRMS was controlled by Thermo Scientific Xcalibur software. Data acquisition and processing for the HPLC-MS was controlled by Waters MassLynx software. (c) Chromatography: flash chromatography was performed using a Teledyne ISCO CombiFlash system. For semi-preparative HPLC Agilent Zorbax Eclipse XDB-C18 or -C8 columns (25 cm x 10 mm, 5 µm particle diameter) were used. An Agilent Zorbax RRHD Eclipse XDB-C18 column (2.1 x 100 mm, 1.8 µm particle diameter) was used in the LC-HRMS *A. fumigatus* mutant profiling analysis. An Agilent Zorbax Eclipse XDB-C18 column (4.6 x 250 mm, 5 µm particle diameter) was used in the HPLC-MS *A. fumigatus* mutant profiling analysis.

Chromatographic enrichment of compounds 1, 3, dimethyl-3, 7-8, 11-13, and 18. Methanol extracts derived from 1-2 L of *A. fumigatus* cultures were fractionated using large-scale reverse-phase flash chromatography on a Teledyne ISCO CombiFlash with a Teledyne C18 gold (100 gram) column with acetonitrile (organic phase) and 0.1 % acetic acid in water (aqueous phase) as solvents at a flow rate of 60 mL/min. A linear ramp from 0 % organic to 100 % organic over 30 min was used and fractions containing compounds of interest were collected, evaporated to dryness, and stored in 8 mL glass vials at -20 °C. Fractions containing compounds **1**, **3**, dimethyl-**3**, **7-**8, **11-13,** and **18** were further purified via semi-preparative HPLC using an Agilent XDB C-18 or C-8 column (25 cm x 10 mm, 5 µm particle diameter) acetonitrile (organic phase) and 0.1 % acetic acid in water (aqueous phase) as solvents at a flow rate of 3.6 mL/min. A solvent gradient scheme was used, starting at 5 % organic for 3 min, followed by a linear increase to 100 % organic over 27 min, holding at 100 % organic for 5 min, then decreasing back to 5 % organic for 0.1 min, and holding at 5 % organic for the final 4.9 min, for a total of 40 min.

***In vivo* stable isotope labeling experiments.** Deuterium- and ^{13}C-labeled amino acids were purchased from Cambridge Isotopes Inc. Each amino acid (labeled or unlabeled control) was added to 50 mL (2 mM) of GMM[1] media in 125 mL Erlenmeyer flasks. Each flask was inoculated with OE::*fsqA* or OE::*fsqA*- *fsqF* (1.0 x 10^6 spores mL^{-1}) and grown for 4 days at 37 °C with shaking at 220 rpm. After 4 days, liquid fungal cultures including fungal tissue and media were frozen using a dry ice acetone bath and lyophilized. The lyophilized residues were extracted with MeOH (25 mL) for 1.5 h with vigorous stirring. Extracts were filtered over cotton, evaporated to dryness, and stored in 4 mL vials at -20 °C. Crude extracts were suspended in 0.5 mL of MeOH and

centrifuged to remove insoluble materials, and the supernatant was subjected to HPLC-MS analysis.

Conversion of 3 into dimethyl-3 for NMR spectroscopic analysis. To a solution of a partially purified fumisoquin C (**3**) in 3:2 toluene/MeOH (2 mL) was added (trimethylsilyl)diazomethane (Sigma-Aldrich) (20 L, 0.04 mmol) as a 2.0 M diethyl ether solution. After stirring at room temperature for 30 min, the mixture was quenched with glacial acetic acid (10 L) and dried by rotary evaporation. The mixture was resuspended in MeOH (100 L) and dimethyl-**3** was purified by semi-preparative HPLC (see section "chromatographic enrichment" above for details).

Synthesis of FsqB substrates. All reagents were purchased and used as is from Sigma-Aldrich. All solvents were purchased and used as is from Fischer Scientific. *N*-methyl-3,4-dihydroxy-DL-phenylalanine (**11**). A 100 mL Schlenk flask under argon at room temperature was charged with 3,4-dihydroxy-DL-tyrosine (394 mg, 2.0 mmol), dissolved in 50 mL of a 7:3 mixture of water and MeOH. To this $NaBH_4$ (228 mg, 6.6 mmol) was added, followed by immediate addition of formaldehyde (0.5 mL of a 37 % solution in H_2O, 6.0 mmol). The reaction was stirred for 1 h at room temperature then quenched by addition of glacial acetic acid (10 mL). The solvents were removed by rotary evaporation, and the product isolated by reverse-phase HPLC. See below for full NMR spectroscopic data of (**12**). *N*-methyl-dopamine. This compound was prepared according to the literature procedure[8].

Table C1. LC-HRMS data of reported compounds

Compound	HR-ESI(+/-) observed (*m/z*)	Ion	Calculated ion formula	Calculated *m/z*	Retention time [min]	Yield of compound (per L culture)*
1	281.1137	$[M+H]^+$	$C_{13}H_{17}N_2O_5^+$	281.1132	0.74	~1–5 mg
2	359.0557	$[M-H]^-$	$C_{13}H_{15}N_2O_8S^-$	359.0555	0.81	~1–5 mg
3	289.0471	$[M-H]^-$	$C_{13}H_9N_2O_6^-$	289.0466	4.31	~10–50 mg
4	241.0255	$[M-H]^-$	$C_{12}H_5N_2O_4^-$	241.0255	4.38	
5	227.0462	$[M-H]^-$	$C_{12}H_7N_2O_3^-$	227.0462	6.58	
7	200.0727	$[M-H]^-$	$C_{12}H_{10}NO_2^-$	200.0717	7.46	~0.5–2 mg
8	216.0676	$[M-H]^-$	$C_{12}H_{10}NO_3^-$	216.0666	5.98	~0.5–2 mg
11	210.0772	$[M-H]^-$	$C_{10}H_{12}NO_4^-$	210.0772	1.71	
12	208.0615	$[M-H]^-$	$C_{10}H_{10}NO_4^-$	208.0165	1.38	
18	293.0565	$[M-H]^-$	$C_{16}H_9N_2O_4^-$	293.0568	5.46	~1–5 mg
21	343.0941	$[M-H]^-$	$C_{17}H_{15}N_2O_6^-$	343.0936	4.31	
22	337.0473	$[M-H]^-$	$C_{17}H_9N_2O_6^-$	337.0466	5.13	

*Numbers indicate estimated production of each compound prior to sample treatment, where significant losses are incurred at each chromatographic step.

Table C2. Predicted homologs of *fsqB*, *fsqC*, and *fsqG* that are clustered in other *Aspergilli*.

Protein	Identity* (%) to genes of				Putative function
A. fumigatus 293	*N. fischeri*	*A. parasiticus*	*A. oryzae*	*A. flavus*	
fsqB	474/497 (95%)	234/501 (47%)	232/501 (46%)	228/501 (46%)	FAD binding domain protein
fsqC	338/365 (93%)	157/321 (49%)	160/343 (47%)	158/325 (49%)	methyltransferase
fsqG	557/598 (93%)	329/619 (53%)	326/619 (53%)	330/619 (53%)	cytochrome P450 monooxygenase

*Percentage similarity values of the *A. fumigatus* 293 ORFs to *N. fischeri, A. parasiticus, A. oryzae,* and *A. flavus* clusters are given in parentheses.

Table C3. Amino acid sequence of polyhistidine-tagged FsqB and software-aided[1] identification of tryptic FsqB peptides analyzed with UHPLC-MS/MS.

MGSSHHHHHHSSGLVPRGSHMSIPNSFIIVGSGVFGLSLAYALSLDDRFADKKIILVDRWNFEPPNATGSVHNPAAAN
ADTSRVIRRDYPHGPYASLALEAMKHWRGKFGENNRYVNQRLLFSGEGSSLTTPPKALETVNYIKKAYAISCELTPGG
RDAVQVLDSLDEVRAFLGNTPSHPPHLPVNKDPAARDLRGYVSNDCGWADAGASIEWLRQEVLRLGRVECVVGEVE
SLVYSDDQRAVKGVKLVDGKVLTAELTVIAAGARSSHILGIPKLCDVYSEFVAYIQLTKEEADELRRRQWPILVNCHRG
VFAVGPDHDNCLKFGHFSYSGIVDVLREASIQVPTRPDGWEAQQKYWSDPRFAFGGEVKVSALGDVDDYENPAAQR
ALADYRLFLLELLGPTGLQGVDTLGLDQSDNLLNNIANRPFTRVRKCWYNDTPALDFVVDYHPSYGKTLFVATGGCD
HAFKFLPIIGEKTLALILRNRGDSAVSLPAGVEPSLEELSELWRFPVELLQDN

Peptide	observed	expected	z	Δ
MGSSHHHHHHSSGLVPR	590.2879	633.9679	3	-131.0400
GSHMSIPNSFIIVGSGVFGLSLAYALSLDDRFADKKIILVDR	1075.2218	1075.2200	3	0.0254
WNFEPPNATGSVHNPAAANADTSR	842.0560	842.0509	3	-0.0066
DYPHGPYASLALEAMK	881.9266	881.9198	2	-0.0010
FGENNR	736.3424	736.3373	1	0.0051
LLFSGEGSSLTTPPK	767.4106	767.4043	2	-0.0020
ALETVNYIK	1050.5813	1050.5757	1	-0.0017
AYAISCELTPGGR	697.8393	697.8330	2	-0.0020
DAVQVLDSLDEVR	1458.7411	1458.7362	1	-0.0024
FLGNTPSHPPHLPVNK	457.2478	457.2409	4	-0.0015
GYVSNDCGWADAGASIEWLR	1113.9964	1113.9900	2	-0.0018
VECVVGEVESLVYSDDQR	1041.9851	1041.9788	2	-0.0020
VLTAELTVIAAGAR	462.2763	462.2695	3	-0.0014
SSHILGIPK	951.5615	951.5549	1	-0.0007
LCDVYSEFVAYIQLTK	974.9887	974.9826	2	-0.0023
EEADELR	431.2024	431.1938	2	0.0027
QWPILVNCHR	661.8428	661.8357	2	-0.0003
GVFAVGPDHDNCLK	764.8631	764.8570	2	-0.0024
FGHFSYSGIVDVLR	798.9108	798.9048	2	-0.0026
EASIQVPTRPDGWEAQQK	1020.5093	1020.5036	2	-0.0032
YWSDPR	412.1909	412.1903	2	0.0012
FAFGGEVK	854.4397	854.4334	1	-0.0009
VSALGDVDDYENPAAQR	910.4254	910.4192	2	-0.0022
LFLLELLGPTGLQGVDTLGLDQSDNLLNNIANRPFTR	1346.7254	1346.7195	3	-0.0042
KCWYNDTPALDFVVDYHPSYGK	1134.4530	892.7407	3	725.1369
TLFVATGGCDHAFK	762.3686	762.3620	2	-0.0013
FLPIIGEK	916.5490	916.5429	1	-0.0012
TLALILR	799.5390	799.5327	1	-0.0009
GDSAVSLPAGVEPSLEELSELWR	1221.1133	1221.1061	2	-0.0002
FPVELLQDN	1074.5451	1074.5393	1	-0.0015
KC(FAD)WYNDTPALDFVVDYHPSYGK	1134.4530	1134.4454	3	0.0009

Table C4. Fungal strains and plasmids used in this study

Strain/plasmid	Description	Reference
***Aspergillus fumigatus* 293 background strains**		
Af293	Wild type	10
Af293.1	*pyrG1*	10
Af293.6	*pyrG1, argB1*	10
TJES1.18	*A.fargB::gpdA(p):: fsqA, pyrG1*	This study
TJES2.20	*ΔfsqA::A.ppyrG, argB1*	This study
TJES3.1	*A.fargB::gpdA(p):: fsqA, A.ppyrG*	This study
TJES4.10	*A.fargB::gpdA(p):: fsqA, ΔfsqF::A.ppyrG*	This study
TJES8.2	*ΔfsqA::A.ppyrG, A.fargB*	This study
TJES13.61	*A.fargB::gpdA(p):: fsqA, ΔfsqB::A.ppyrG*	This study
TJES14.68	*A.fargB::gpdA(p):: fsqA, ΔfsqC::A.ppyrG*	This study
TJES15.91	*A.fargB::gpdA(p):: fsqA, ΔfsqG::A.ppyrG*	This study
***Aspergillus flavus* NRRL3357 background strains**		
NRRL3357.5	*pyrG1*	11
TJW149.27	*Δku70::pyrG*	This study
TJES19.1	*ΔpyrG, Δku70*	This study
TJES 23.3	*A.ppyrG::gpdA(p)::imqA, Δku70*	This study
TJES 27.1	*ΔimqA::A.ppyrG, Δku70*	This study
Plasmids		
pJMP4	*A. fumigatus argB*	12
pJMP8.1	*A. nidulans* gpdA promoter (truncated to 1.5 kb) in pBluescript	13
pJMP9.1	*A. nidulans gpdA(p) + A parasiticus pyrG*	14
pJES1.2	*gpdA(p):: fsqA* in *pJMP8.1*	This study
pJES2.7	*A. fumigatus argB::gpdA(p):: fsqA*	This study
pJES13.2	C-terminal 6His tagged *fsqF* A-domain in pET30a vector	This study
pJW24	*A parasiticus pyrG* in pBluescript	15

pXX = plasmid, TXX = original transformant

Table C5. PCR primer sets used in this study

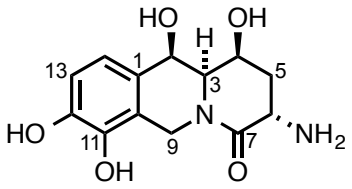
Name of the primer	Oligonucleotide sequence (5'-3')	Uses
Afu6g03380_FOR_Nprobe	CTTTCACGTGGACACTGCGC	Northern probe
Afu6g03380_REV_Nprobe	TACTGCTCCAACCAGCACCG	Northern probe
Afu6g03390_FOR_Nprobe	TCAGGGAGATATGGTGGCG	Northern probe
Afu6g03390_REV_Nprobe	CAACGCAGCAGGTAGTCACG	Northern probe
Afu6g03400_FOR_Nprobe	CTTCCAAGCCCAACAAGCC	Northern probe
Afu6g03400_REV_Nprobe	AATCTCGTAGGCCTCCAGCG	Northern probe
Afu6g03430_FOR_Nprobe	TGGCCTATCACACCAGTGGC	Northern probe
Afu6g03430_REV_Nprobe	GTGCAGCCTGAATCTCACGG	Northern probe
Afu6g03420_FOR_Nprobe	GAGCTCGACAATGGTGAGCG	Northern probe
Afu6g03420_REV_Nprobe	CAATCACATCGGCATGCGG	Northern probe
Afu6g03440_FOR_Nprobe	GATGGAACTTCGAGCCACCC	Northern probe
Afu6g03440_REV_Nprobe	GGCGATAATCTGCCAACGCC	Northern probe
Afu6g03450_FOR_Nprobe	CTACATCGCCTGCGATGTGG	Northern probe
Afu6g03450_REV_Nprobe	TAGCACACGCGCCAGATACC	Northern probe
Afu6g03460_FOR_Nprobe	GCGACTTCGCGACTCGGAAT	Northern probe
Afu6g03460_REV_Nprobe	CCATCACAAACTCGGTCCCG	Northern probe
Afu6g03470_FOR_Nprobe	AACTGCGCTCCAAAACCGCC	Northern probe
Afu6g03470_REV_Nprobe	ATCCACAAGGGCGATCTGGC	Northern probe
Afu6g03490_FOR_Nprobe	GTTCTCAGGGGATGTGACCG	Northern probe
Afu6g03490_REV_Nprobe	ACAAGTTCGCCCTTCGCTCCG	Northern probe
Afu6g03500_FOR_Nprobe	GGTGCTCAAGGAACAGAGGG	Northern probe
Afu6g03500_REV_Nprobe	GCCAAGAGGTCATTCTGCCC	Northern probe
Afu6g03510_FOR_Nprobe	AACACCCGATACCAGCTCGC	Northern probe
Afu6g03510_REV_Nprobe	ATGGGCCACCCATTGATGGC	Northern probe
Afu6g03520_FOR_Nprobe	ATGCCATCATCACCGGTGCC	Northern probe
Afu6g03520_REV_Nprobe	CGAGCATGGACAATAGCC	Northern probe
A.ppyrG_T7 FOR	CGTAATACGACTCACTATAGGG	Amplification of A.ppyrG from pJMP9.1
A.ppyrGR_Rev	ATTCGACAATCGGAGAGGCTGC	Amplification of A.ppyrG from pJMP9.1
Afu6g03430_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCG AATGCTTCAGCTGGAGTGTCTCC	Amplification of *fsqA* 3' flanking region
Afu6g03430_3'R_flank	TACAGCGACGACCAACGAGC	Amplification of *fsqA* 3' flanking region
Afu6g03430_5'F_flank	TAAGAGCGGAGACTGGTGGC	Amplification of *fsqA* 5' flanking region
Afu6g03430_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATT ACGTCTGCAAGGGTTTACGAGGG	Amplification of *fsqA* 5' flanking region
Afu6g03440_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCG AATTCGGAACTCTGGAGGTTCCC	Amplification of *fsqB* 3' flanking region
Afu6g03440_3'R_flank	ACTGCGCGACAAATGCAGCC	Amplification of *fsqB* 3' flanking region
Afu6g03440_5'F_flank	GTCTCGTCACTTACCCTGCC	Amplification of *fsqB* 5'

		flanking region
Afu6g03440_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATTACGAAAGAGACAGCCGGGATCCG	Amplification of *fsqB* 5' flanking region
Afu6g03450_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCGAATCTTGCTGCGGAAATCGAGCG	Amplification of *fsqC* 3' flanking region
Afu6g03450_3'R_flank	CACGGTAAAAGCCCAGTCCG	Amplification of *fsqC* 3' flanking region
Afu6g03450_5'F_flank	GATGTAGGCCACGAACTCGC	Amplification of *fsqC* 5' flanking region
Afu6g03450_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATTACGAGGATGCCAAAAGCCCACCG	Amplification of *fsqC* 5' flanking region
Afu6g03480_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCGAATCGCGGGCATCTAGTATTCGG	Amplification of *fsqF* 3' flanking region
Afu6g03480_3'R_flank	ACTTGCGCAACCAGCTGTGC	Amplification of *fsqF* 3' flanking region
Afu6g03480_5'F_flank	GAATCTGAGCGCTTGTCGCG	Amplification of *fsqF* 5' flanking region
Afu6g03480_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATTACGAAGAAAGGCGAAACGGAGCG	Amplification of *fsqF* 5' flanking region
Afu6g03490_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCGAATATGGACTCCAGTCAGGACCG	Amplification of *fsqG* 3' flanking region
Afu6g03490_3'R_flank	ATCCACCTCGTGGAGAAGCC	Amplification of *fsqG* 3' flanking region
Afu6g03490_5'F_flank	TGTGTCACGAAGGCAGTGCG	Amplification of *fsqG* 5' flanking region
Afu6g03490_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATTACGGTGCCCATCGTCCAATACGG	Amplification of *fsqG* 5' flanking region
fqsF_Adomain_xp_5'F	AATAA**GCGGCCGC**ACCTGGAACACCGTGGTTGC	Amplification of *fsqF* A-domain **(NotI cut site)**
fqsF_Adomain_xp_3' R	ATTAGC**CTCGAG**TTCGTTGCCCCGAGTGTGC	Amplification of *fsqF* A-domain **(XhoI cut site)**
fsqD_xp_5' F	GTAGGCTA**GAATTC**TCACACCAGCCTCTCTTGGC	Amplification of *fsqD* ORF **(EcoRI cut site)**
fsqD_xp_3' R	CATTAT**AAGCTT**GAGAAAGGAGTAGCGGACCTCTTCCC	Amplification of *fsqD* ORF **(HindIII cut site)**
Afu6g03430_NcoI_FOR	CAGATG**CCATGG**ACGACAAGCATGGCC	Amplification of *fsqA* ORF **(NcoI cut site)**
Afu6g03430_NotI_REV	GCAACC**GCGGCCGC**AGACAGCGCGGTATCACTG	Amplification of *fsqA* ORF **(NotI cut site)**
flvku70F5	ACATCTCTTCCGTCAAAGGCGC	Amplification of *ku70* 5’ flanking region
flavku70R5	CGATATCAAGCTATCGATACCTCGACTCTGTGTTGAGAGTCGTAAGTCATGAATTGCG	Amplification of *ku70* 5’ flanking region
flvku70F3	GTCGCTGCAGCCTCTCCGATTGTCGAATGACAACGCTAGTATTGGTTACGAGAG	Amplification of *ku70* 3’ flanking region

	ACAG	
flavku70R3	AGAATGGCTACGTCAACCTCCG	Amplification of *ku70* 3’ flanking region
Fku70IF	ATGAGGAAGAGGAGGAGACCG	Amplification of Δ*ku70* cassette
Fku70IR	CACTTTTCAATCGTGCGAGCCG	Amplification of Δ*ku70* cassette
pet28_fsqB_3’_fwd	TGGTGCCGCGCGGCAGCCATATGTC TATCCCTAACTCTTTCATCATTGT	Amplification of fsqB, pET28b+ 3’ destination
fsqB_5’_pet28_rev	CTCAGCTTCCTTTCGGGCTTTGTTACTA GTTGTCCTGTAGTAGTTCCACGG	Amplification of fsqB pET28b+ 5’ destination

Table C6. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for fumisoquin A, 1, in methanol-d_4.

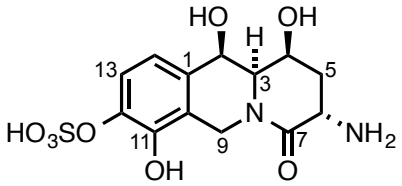
Chemical shifts were referenced to δ(C<u>H</u>D_2OD) = 3.31 ppm and δ(^{13}<u>C</u>HD_2OD) = 49.0 ppm.^{13}C chemical shifts were determined via HMBC and HSQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. NOESY correlations were observed using a mixing time of 400 ms. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC	NOESY
1	128.03				
2	70.99	2-H	4.94 ($J_{2,3}$ = 1.9)	1, 3, 4, 7, 10, 11, 14	3, 4, 14
3	60.44	3-H	3.73 ($J_{3,2}$ = 1.9, $J_{3,4}$ = 8.7)	1, 7	2, 4, 5b
4	67.19	4-H	4.65 ($J_{4,3}$ = 8.7, $J_{4,5a}$ = 7.6, $J_{4,5b}$ < 1)	5	2, 5a, 5b
5	33.67	5-H_a	2.46 ($J_{5a,4}$ = 7.6, $J_{5a,5b}$ = 13.0, $J_{5a,6}$ = 5.7)	3, 4, 6, 7	4, 6, 5b
		5-H_b	2.06 ($J_{5b,4}$ < 1, $J_{5b,5a}$ = 13.0, $J_{5b,6}$ = 12.4)	3, 4, 6, 7	3, 4, 5a
6	47.88	6-H	4.29 ($J_{6,5a}$ = 5.7, $J_{6,5b}$ = 12.4)	5, 7	5a
7	169.75				
8					
9	42.35	9-H_a	4.66 ($J_{9Ha,9Hb}$ = 18.0)	1, 2, 3, 6, 7, 10, 11, 12, 13	
		9-H_b	4.78 ($J_{9Hb,9Ha}$ = 18.0)	1, 2, 3, 6, 7, 10, 11, 12, 13	3
10	119.92				
11	142.82				
12	146.06				
13	114.23	13-H	6.73 ($J_{13,14}$ = 8.0)	1, 10, 11, 12	14
14	120.04	14-H	6.70 ($J_{14,13}$ = 8.0)	1, 2, 4, 9, 10, 12, 13	2,1 3

Table C7. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for fumisoquin B, 2, in methanol-d_4.

Chemical shifts were referenced to δ(C<u>H</u>D$_2$OD) = 3.31 ppm and δ(^{13}<u>C</u>HD$_2$OD) = 49.0 ppm.^{13}C chemical shifts were determined via HMBC and HSQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. ROESY correlations were observed using a mixing time of 500 ms. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

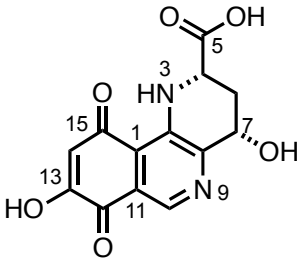


No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC[a]	ROESY
1	134.03				
2	70.75	2-H	4.98 ($J_{2,3}$ = 2.0)	1, 3, 10	4, 14
3	60.19	3-H	3.71 ($J_{3,2}$ = 2.0, $J_{3,4}$ = 9.2)		5a
4	67.30	4-H	4.62 ($J_{4,3}$ = 9.2, $J_{4,5a}$ = 6.9, $J_{4,5b}$ = 7.9)	6	2
5	34.69	5-H$_a$	1.94 ($J_{5a,4}$ = 6.9, $J_{5a,5b}$ = 13.0, $J_{5a,6}$ = 11.5)	4, 6, 7	3
		5-H$_b$	2.43 ($J_{5b,4}$ = 7.9, $J_{5b,5a}$ = 13.0, $J_{5b,6}$ = 5.8)	4, 6, 7	
6	47.88	6-H	4.02 ($J_{6,5a}$ =11.5, $J_{6,Hb}$ = 5.8)	5, 7	
7	171.61				
8					
9	42.26	9-H$_a$	4.63 ($J_{9Ha,9Hb}$ = 18.5)	3	
		9-H$_b$	4.81 ($J_{9Hb,9Ha}$ = 18.5)	10, 11	
10	121.61				
11	147.05				
12	140.92				
13	122.08	13-H	7.23 ($J_{13,14}$ = 8.1)	1, 11, 12w	
14	119.96	14-H	6.83 ($J_{14,13}$ = 8.1)	2, 10, 12, 13	2

[a]w: weak correlation (less than ~10% of the intensity of strongest signal)

Table C8. ^{1}H and ^{13}C spectroscopic data for fumisoquin C, 3, in DMSO-d_6.

Chemical shifts were referenced to δ($C\underline{H}D_2SOCD_3$) = 2.50 ppm and δ($^{13}\underline{C}HD_2SOCD_3$) = 39.52 ppm.$^{13}C$ chemical shifts were determined via HMBC and HSQC spectra. Spectra were acquired using the Varian INOVA 600 spectrometer, except for the HSQC, which was acquired using the Bruker Avance 800 spectrometer. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectrum. NOESY correlations were observed using a mixing time of 600 ms. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

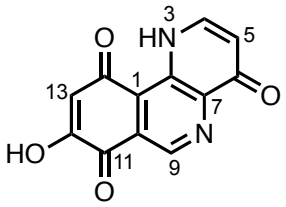


No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC	NOESY
1	115.62				
2	138.33				
3		3-NH	9.46 ($J_{3,4}$ = 1.5)	1, 4, 5, 6, 8	4
4	53.27	4-H	3.73 ($J_{4,3}$ = 1.5, $J_{4,6a}$ =9.6, $J_{4,6b}$ = 5.4)	2, 5, 6, 7	3, 7
5	173.40				
6	32.31	6-Ha	1.69 ($J_{6a,4}$ =12, $J_{6a,6b}$ = 12.8, $J_{6a,7}$ = 9.5)	4, 5, 7, 8	
		6-Hb	2.38 ($J_{6b,4}$ =12, $J_{6a,6b}$ = 12.8, $J_{6b,7}$ = 5.3)	4, 5, 7, 8	
7	66.85	7-H	4.57 ($J_{7,6a}$ = 9.5, $J_{7,6b}$ = 5.3)	2, 4, 6, 8	4
8	152.69				
9					
10	130.73	10-H	7.98	1, 8, 11, 12, 13, 15	
11	124.93				
12	188.04				
13[a]	184.60				
14	107.92	14-H	5.24	1, 12, 13, 15	
15[a]	170.85				

[a]Position 13 and 15 carbon chemical shifts are interchangeable

Table C9. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for compound 4 in DMSO-d_6.

Chemical shifts were referenced to δ($C\underline{H}D_2SOCD_3$) = 2.50 ppm and δ($^{13}\underline{C}HD_2SOCD_3$) = 39.52 ppm.$^{13}C$ chemical shifts were determined via HMBC and HSQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectrum. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

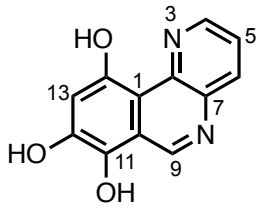


No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC[a]
1	127.41			
2	134.33			
3		3-NH	13.71 ($J_{3,4}$ = 6.6, $J_{3,5}$ = 1.7)	
4	140.30	4-H	8.06 ($J_{4,3}$ = 6.6, $J_{4,5}$ = 7.3)	2, 5, 6
5	111.88	5-H	6.24 ($J_{5,4}$ = 7.3, $J_{5,3}$ = 1.7)	4, 7
6	175.28			
7	144.61			
8				
9	141.85	9-H	8.91	1, 2w, 7, 10, 11w, 12w
10	125.00			
11	185.04			
12	180.67			
13[b]	108.10	13-H	5.45	1, 2w, 11, 12w, 14w
14[b]	172.04			

[a]w: weak correlation (less than ~10% of the intensity of strongest signals); [b]Position 12 and 14 carbon chemical shifts are interchangeable

Table C9. ^{1}H (800 MHz) and or ^{13}C (200 MHz) NMR spectroscopic data for compound 5 in DMSO-d_6.

Chemical shifts were referenced to δ($C\underline{H}D_2SOCD_3$) = 2.50 ppm and δ($^{13}\underline{C}HD_2SOCD_3$) = 39.52 ppm.^{13}C chemical shifts were determined via HMBC and HSQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectrum. NOESY correlations were observed using a mixing time of 600 ms. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

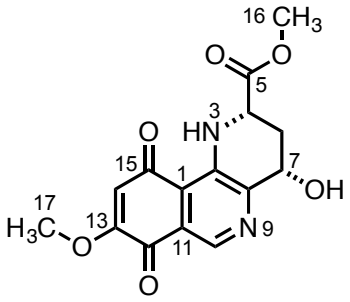


No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC[a]	NOESY
1	108.19				
2	142.74				
3					
4	146.12	4-H	8.87 ($J_{4,5}$ = 4.7)	2, 5, 6	5, 14-OH
5	121.77	5-H	7.75 ($J_{5,4}$ = 4.7, $J_{5,6}$ = 8.3)	4, 7	4, 6
6	136.97	6-H	8.48 ($J_{6,5}$ = 8.3)	2, 4	5
7	136.52				
8					
9	149.94	9-H	9.52	1, 2w, 7, 10, 11, 14w	
10	118.41				
11	133.63				
12	146.74				
13	106.91	13-H	6.97	1, 2w, 11, 12, 14	
14	150.40				
14		14-OH	13.42	1, 12, 13, 14	4

[a]w: weak correlation (less than ~10% of the intensity of strongest signal)

Table C10. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for dimethyl-fumisoquin C (dimethyl-3) in DMSO-d_6.

Chemical shifts were referenced to δ($C\underline{H}D_2SOCD_3$) = 2.50 ppm and δ($^{13}\underline{C}HD_2SOCD_3$) = 39.52 ppm.^{13}C chemical shifts were determined via HMBC and HSQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectrum. ROESY correlations were observed using a mixing time of 600 ms. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

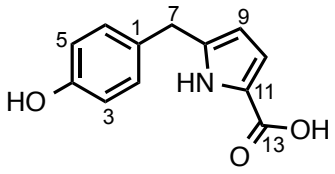


No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC[a]	ROESY
1	112.56				
2	138.34				
3		3-NH	9.22 ($J_{3,4}$ = 3.8)	1,2,4,6,8	4
4	49.12	4-H	4.49 ($J_{4,3}$ = 3.8, $J_{4,6a}$ = 6.3) ($J_{4,6b}$ = 12.2)	2,5,6,7,8	3, 6a, 6b, 7
5	172.07				
6	30.13	6-Ha	2.24 ($J_{6a,4}$ =6.3, $J_{6a,6b}$ = 13.5, $J_{6a,7}$ = 3.6)	4,5,7,8	4, 6b, 7
		6-Hb	2.38 ($J_{6b,4}$ =12.2, $J_{6a,6b}$ = 13.5) ($J_{6b,7}$ = 9.0)	4,5,7,8	4, 6a, 7
7	65.09	7-H	4.69 ($J_{7,6a}$ = 3.6, $J_{7,6b}$ = 9.0)		4, 6a, 6b, 7OH
7		7-OH	5.61 ($J_{7OH,7}$ = 4.2)	2,4,8	7
8	153.51				
9					
10	132.52	10-H	8.31	1,2w,8,11,12,13,15	
11	123.43				
12	179.63				
13	159.07				
14	111.20	14-H	6.32	1,2,12,13,15	17
15	187.38				
16	51.58	16-H3	3.63	5	4
17	56.24	17-H3	3.86	13,14	14

[a]w: weak correlation (less than ~10% of the intensity of strongest signal)

Table C11. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for compound 8 in methanol-d_4.

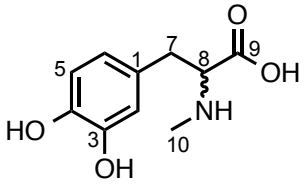
Chemical shifts were referenced to δ(C<u>H</u>D_2OD) = 3.31 ppm and δ(^{13}<u>C</u>HD_2OD) = 49.0 ppm.^{13}C chemical shifts were determined via an HMBC spectrum. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectrum. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC
1				
2	129.62	2-H	7.03 ($J_{2,3}$ = 8.5)	1, 3, 4, 7
3	130.63	3-H	6.70 ($J_{3,2}$ = 8.5)	1, 4, 5
4	155.91			
5	130.63	5-H	6.70 ($J_{5,6}$ = 8.5)	1, 4, 5
6	129.62	6-H	7.03 ($J_{6,5}$ = 8.5)	1, 3, 4, 7
7	33.04	7-H_2	3.83	2, 3, 5, 6, 8, 9
8	138.30			
9	108.30	9-H	5.80 ($J_{9,10}$ = 3.5)	7, 8, 10, 11, 13
10	116.12	10-H	6.62 ($J_{10,9}$ = 3.5)	8, 9, 11, 13
11	122.63			
12		12-NH		
13	164.53			

Table C12. ^{1}H (500 MHz) and ^{13}C (125 MHz) NMR spectroscopic data for compound 11 in water-d_2.

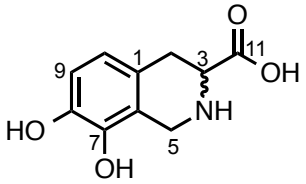
Chemical shifts were referenced to δ(HOD) = 2.79 ppm and δ(^{13}CH$_3$OD) = 49.0 ppm.^{13}C chemical shifts were determined via an HMBC and HSQC spectrum. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectrum. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC
1	127.21			
2	117.10	2-H	6.84 ($J_{2,6}$ = 1.8)	4, 6, 7
3	144.36			
4	143.63			
5	116.49	5-H	6.92 ($J_{5,6}$ = 8.0)	1, 3
6	121.86	6-H	6.76 ($J_{6,5}$ = 8.0, $J_{6,2}$ = 1.8)	2, 4, 5, 7
7	35.06	7-H$_2$	3.14 ($J_{7,8}$ = 6.0)	1, 2, 6, 8, 9
8	64.71	8-H	3.82 ($J_{8,7}$ = 6.0)	1, 7, 9, 10
9	173.12			
10	32.16	10-H$_3$	2.71	8

Table C12. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for compound 12 in water-d_2.

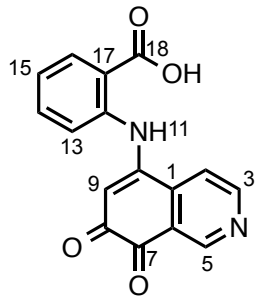
Chemical shifts were referenced to δ(HOD) = 2.79 ppm and δ(^{13}CH$_3$OD) = 49.0 ppm.^{13}C chemical shifts were determined via an HMBC and HSQC spectrum. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectrum. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC
1	124.11			
2	28.50	2-Ha	3.08 ($J_{2a,2b}$ = 16.0, $J_{2a,3}$ = 11.5)	1, 3, 6, 10, 11
		2-Hb	3.32 ($J_{2b,2a}$ = 16.0, $J_{2b,3}$ = 5.0)	1, 3, 6, 10, 11
3	56.25	3-H	3.98 ($J_{3,2a}$ = 11.5, $J_{3,2b}$ = 5.0)	1, 2, 5, 11
4		4-NH		
5	40.60	5-Ha	4.25 ($J_{5a,5b}$ = 16.0)	1, 3, 6, 7
		5-Hb	4.54 ($J_{5b,5a}$ = 16.0)	1, 3, 6, 7
6	116.69			
7	140.97			
8	142.61			
9	116.49	9-H	6.93 ($J_{9,10}$ = 8.4)	1, 7
10	121.86	10-H	6.80 ($J_{10,9}$ = 8.4)	2, 6, 8
11	173.12			

Table C13. ^{1}H (600 MHz) and ^{13}C (151 MHz) NMR spectroscopic data for compound 18 in DMSO-d_6.

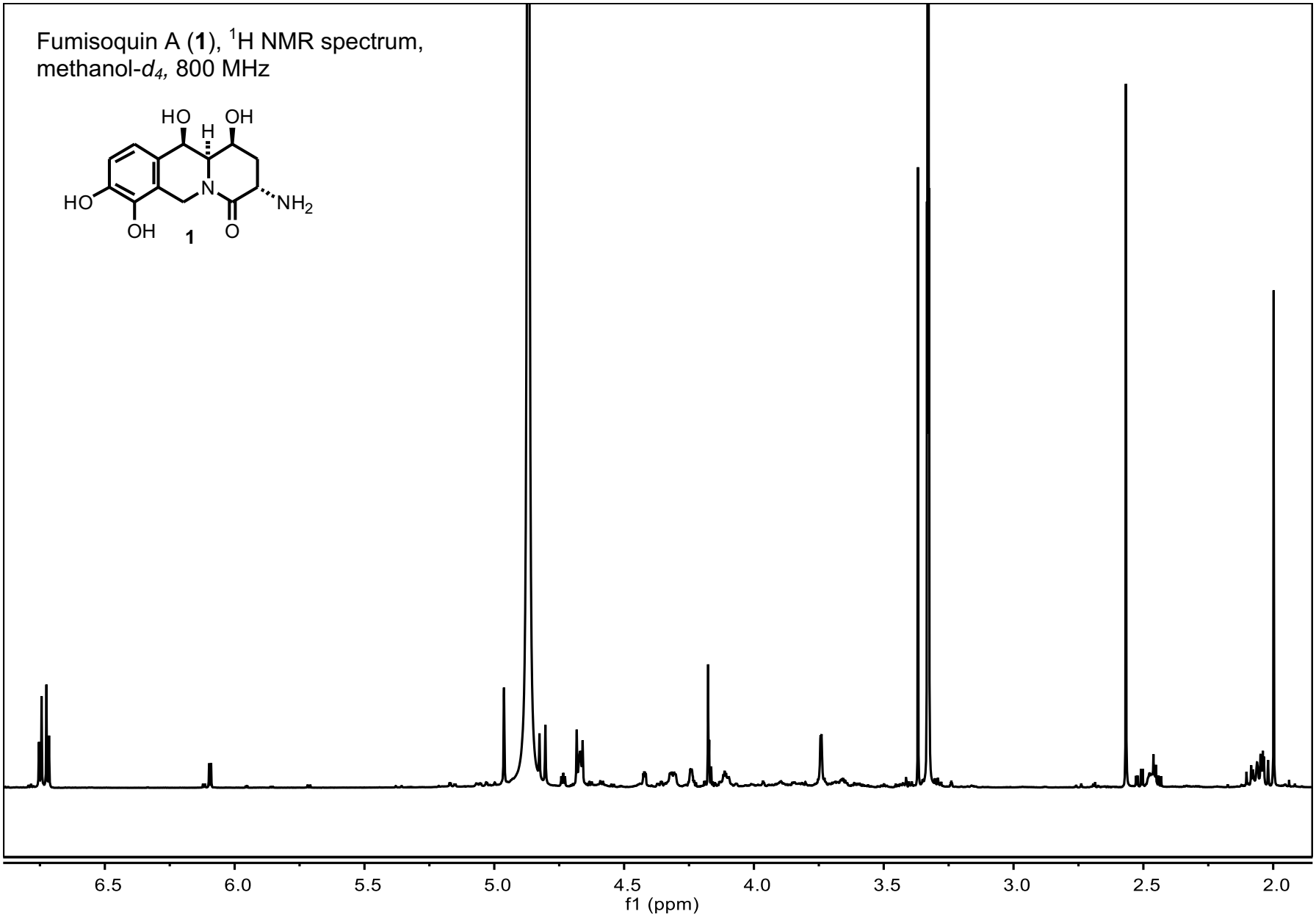
Chemical shifts were referenced to δ(C<u>H</u>D_2SOCD_3) = 2.50 ppm and δ(^{13}<u>C</u>HD_2SOCD_3) = 39.52.^{13}C chemical shifts were determined via HMBC and HSQC spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H spectrum. NOESY correlations were observed using a mixing time of 600 ms. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

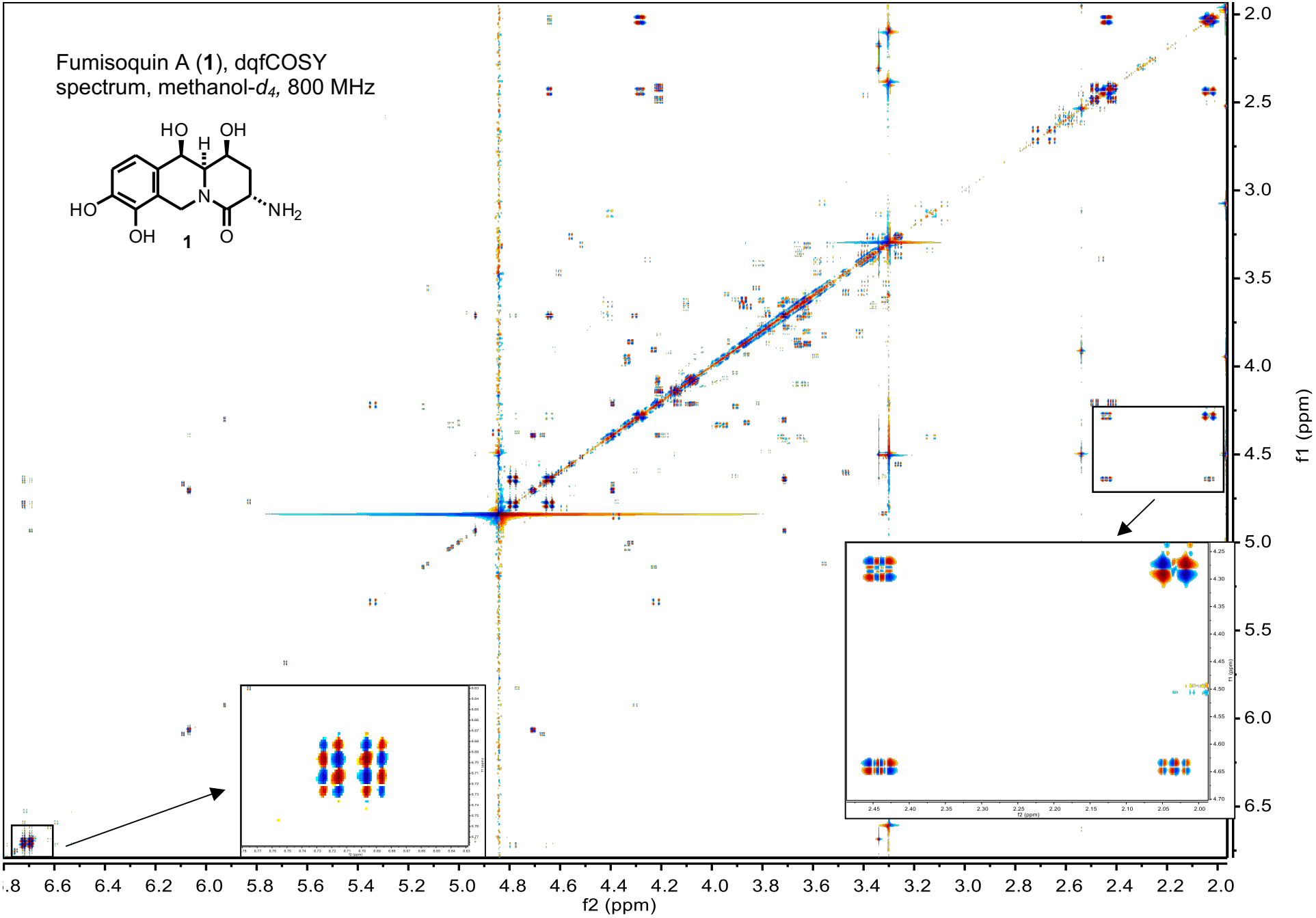


No.	δ_c	Proton	δH (J_{HH}[Hz])	HMBC[a]	NOESY
1	140.40				
2	117.23	2-H	8.05 ($J_{2,3}$ = 5.2)	3, 6, 7, 10	
3	154.59	3-H	9.02 ($J_{3,2}$ = 5.2)	1, 2, 5	
4					
5	148.00	5-H	9.08 ($J_{5,2}$ = 2.3)	1, 3, 6, 7	
6	125.13				
7	181.43				
8	175.31				
9	102.40	9-H	6.45	1, 7, 8, 10w	13
10	149.45				
11		11-NH			
12	142.36				
13	120.03	13-H	7.48 ($J_{13,14}$ = 8.5)	15, 17	9
14	130.23	14-H	7.45 ($J_{14,13}$ = 8.5, $J_{14,15}$ = 7.5)	12, 16	
15	122.68	15-H	7.11 ($J_{15,14}$ = 7.5, $J_{15,16}$ = 7.5)	13, 17	
16	131.10	16-H	8.01 ($J_{16,15}$ = 7.5)	12, 14, 18	
17	126.86				
18	168.47				

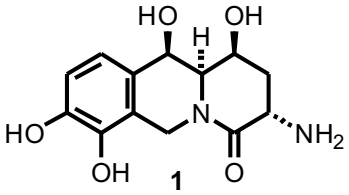
[a]w: weak correlation (less than ~10% of the intensity of strongest signal

NMR spectra of compounds 1-5, dimethyl-3, 8, 11-13, and 18

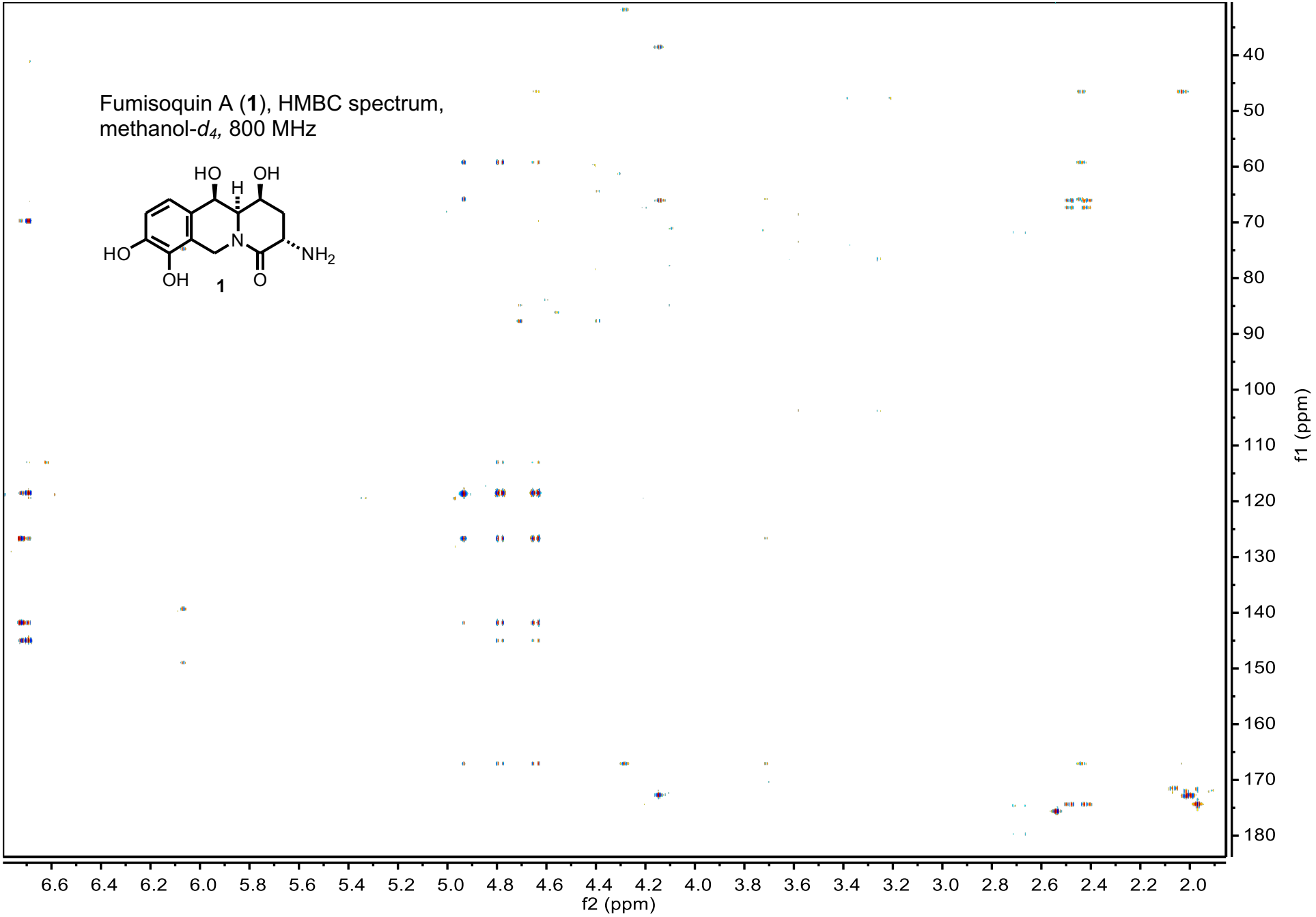


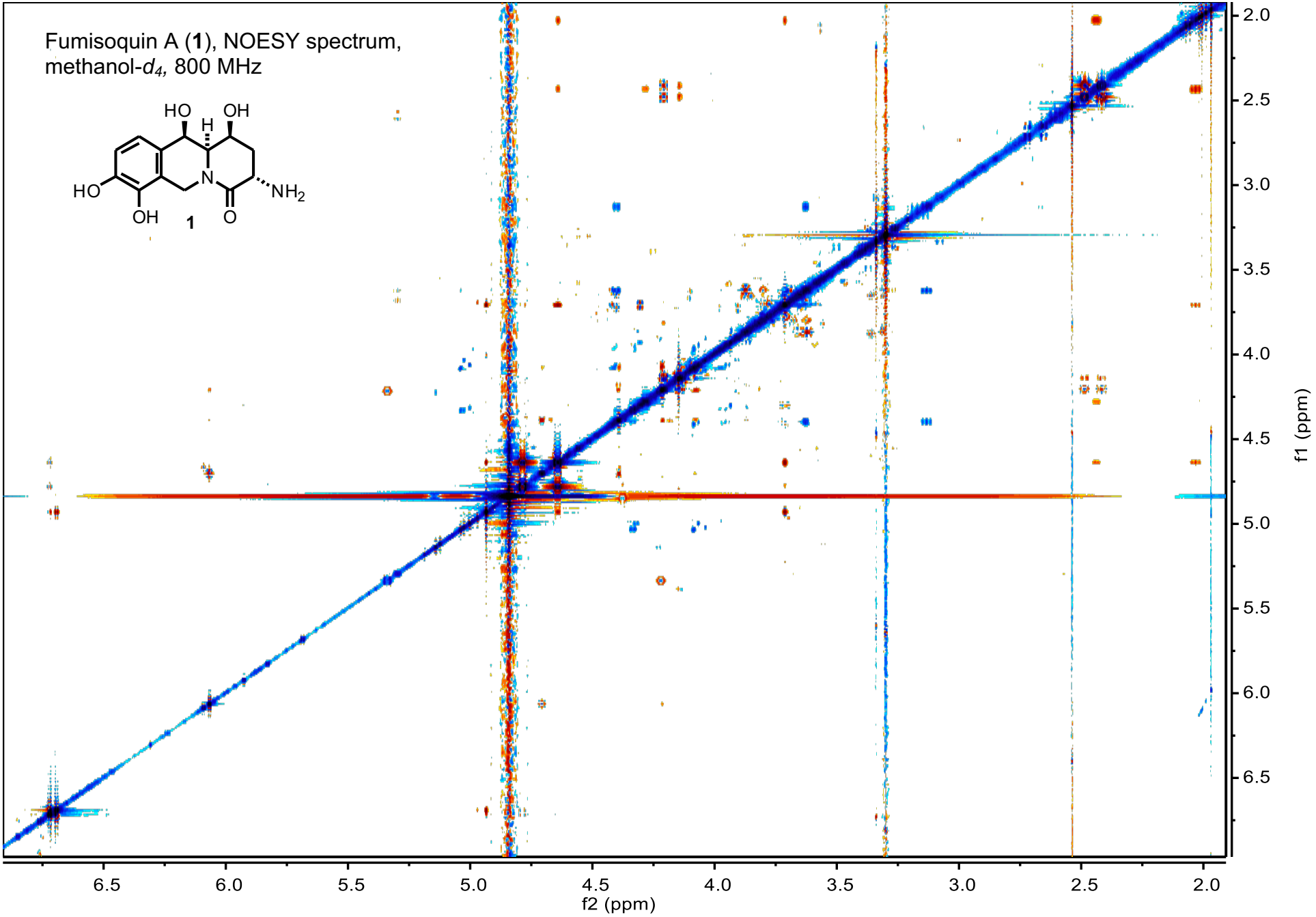

Fumisoquin A (1), dqfCOSY spectrum, methanol-d_4, 800 MHz
HO
OH
H
N
HO
OH
O
NH_2
1
f2 (ppm)
f1 (ppm)

Fumisoquin A (**1**), coupled HSQC spectrum, methanol-d_4, 800 MHz

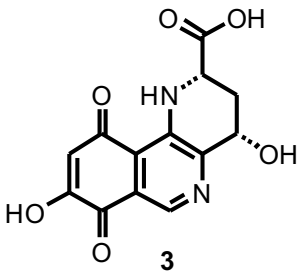


1


Fumisoquin A (1), HMBC spectrum,
methanol-d4, 800 MHz
HO
H
OH
HO
N
NH2
OH
1
O
f2 (ppm)
f1 (ppm)

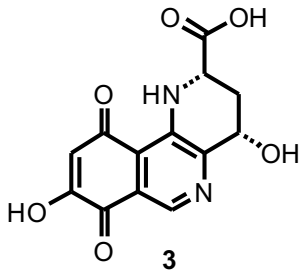

Fumisoquin A (1), NOESY spectrum,
methanol-d4, 800 MHz
HO
OH
H
HO
N
NH2
OH
1
O
f2 (ppm)
f1 (ppm)
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0

Fumisoquin C (**3**), containing about 40% of decomposition product **4,** ^{1}H NMR spectrum, methanol-d_4, 600 MHz



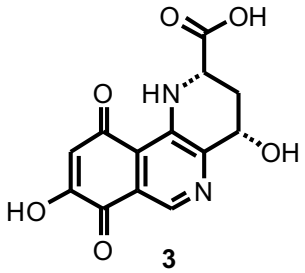
3

Fumisoquin C (**3**), containing about 40% of decomposition product **4,** dqfCOSY spectrum, methanol-d_4, 600 MHz



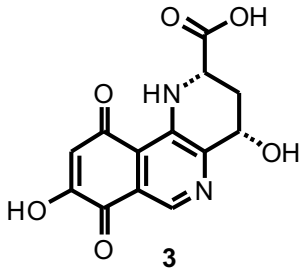
3

Fumisoquin C (**3**), containing about 40% of decomposition product **4,** ^{1}H NMR spectrum, DMSO-d_6, 600 MHz



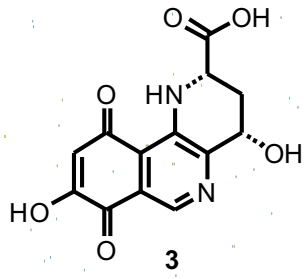
3

Fumisoquin C (**3**), containing about 40% of decomposition product **4**, coupled HSQC spectrum, DMSO-d_6, 800 MHz

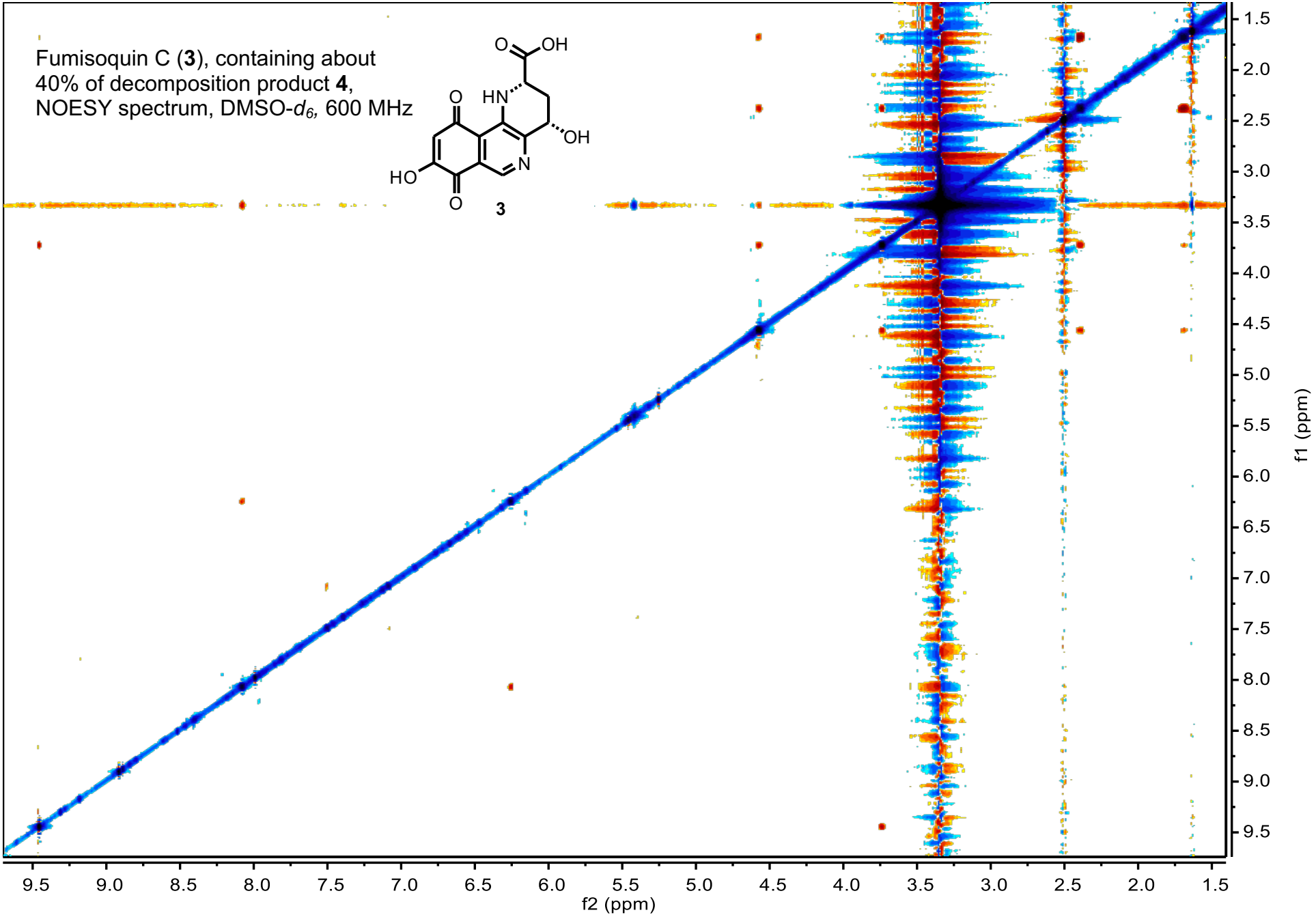


3

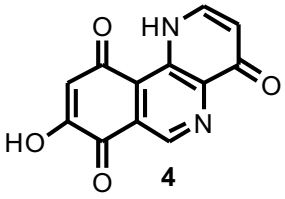
Fumisoquin C (**3**), containing about
40% of decomposition product **4**,
HMBC spectrum, DMSO-d_6, 600 MHz



3

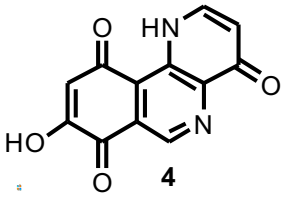

Fumisoquin C (3), containing about
40% of decomposition product 4,
NOESY spectrum, DMSO-d6, 600 MHz
O
OH
HN
O
OH
HO
N
O
3
f2 (ppm)
f1 (ppm)

Compound **4,** ^{1}H NMR spectrum,
DMSO-d_6, 800 MHz



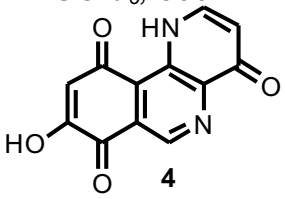
4

Compound **4,** dqfCOSY spectrum,
DMSO-d_6, 800 MHz



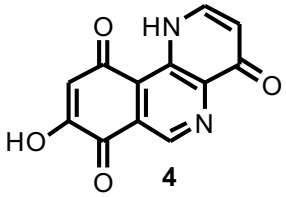
4

Compound **4,** coupled HSQC spectrum,
DMSO-d_6, 800 MHz

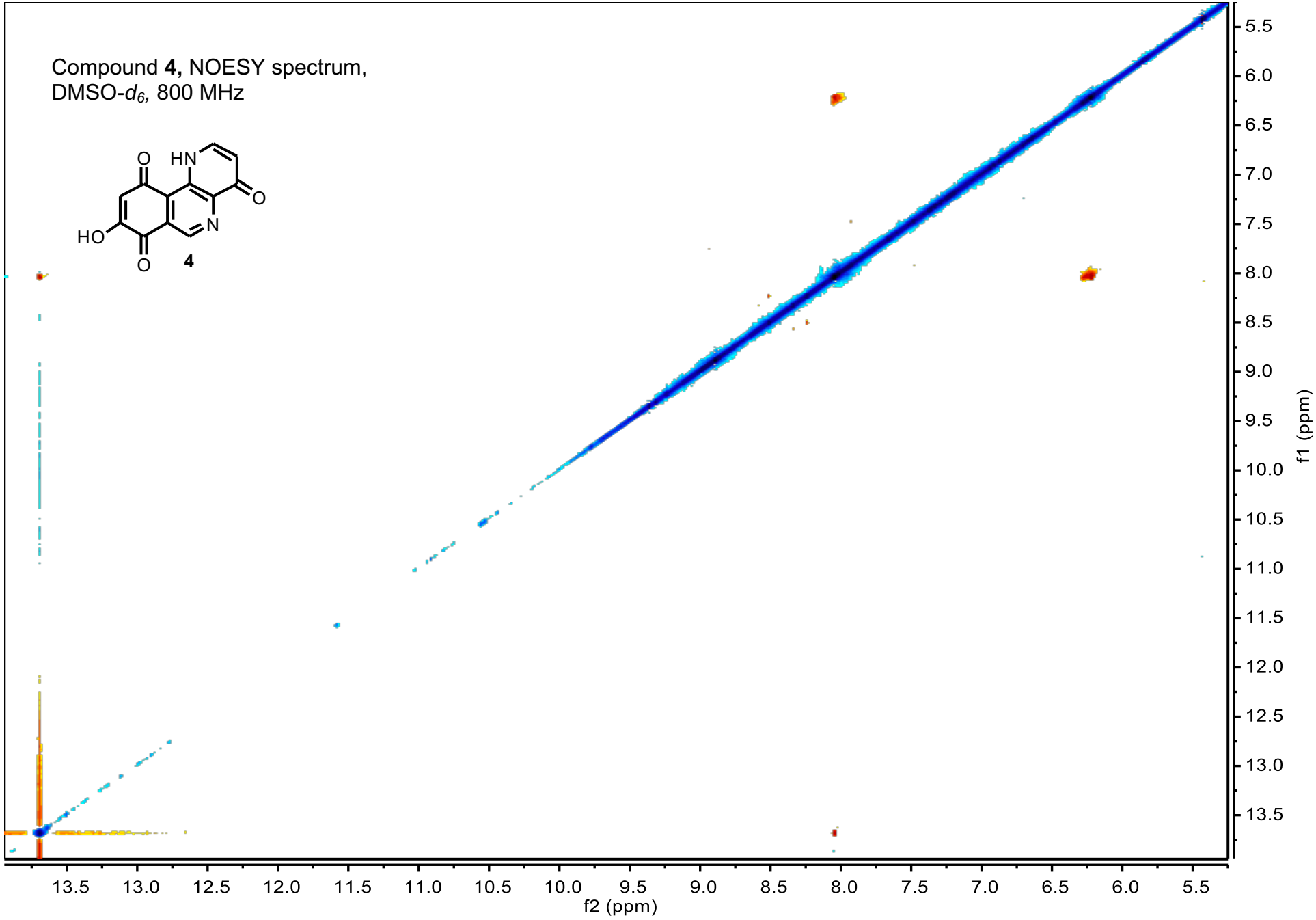


4

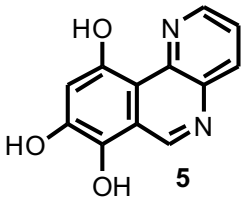
Compound **4,** HMBC spectrum,
DMSO-d_6, 800 MHz



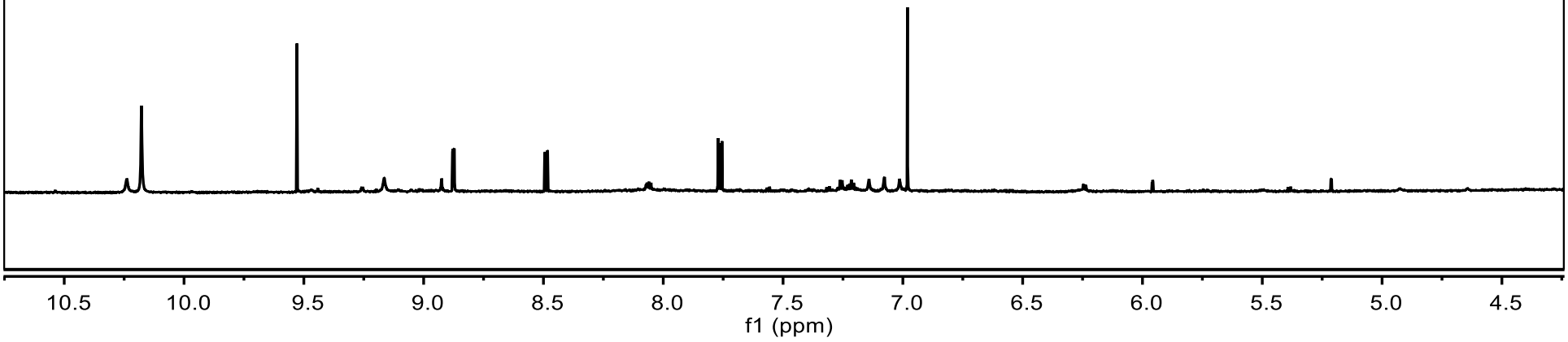
4


Compound **4,** NOESY spectrum,
DMSO-d_6, 800 MHz
O
HN
O
HO
N
O
4
f2 (ppm)
f1 (ppm)
13.5
13.0
12.5
12.0
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5

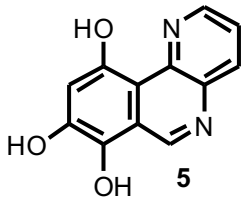
Compound **5,** ^{1}H NMR spectrum,
DMSO-d_6, 800 MHz



5

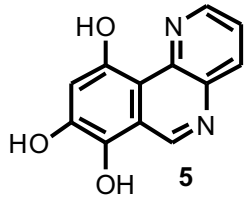


Compound **5,** dqfCOSY spectrum,
DMSO-d_6, 800 MHz



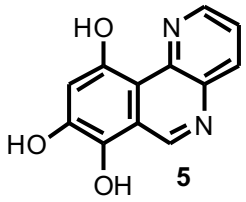
5

Compound **5,** coupled HSQC spectrum,
DMSO-d_6, 800 MHz



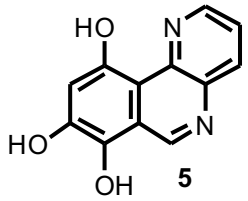
5

Compound **5,** HMBC spectrum,
DMSO-d_6, 800 MHz



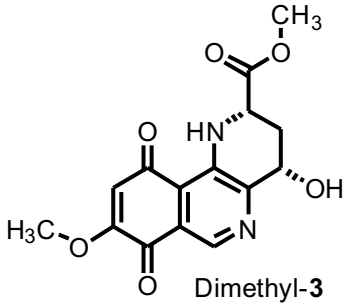
5

Compound **5,** NOESY spectrum
(for confirming exchangeable protons),
DMSO-d_6, 800 MHz

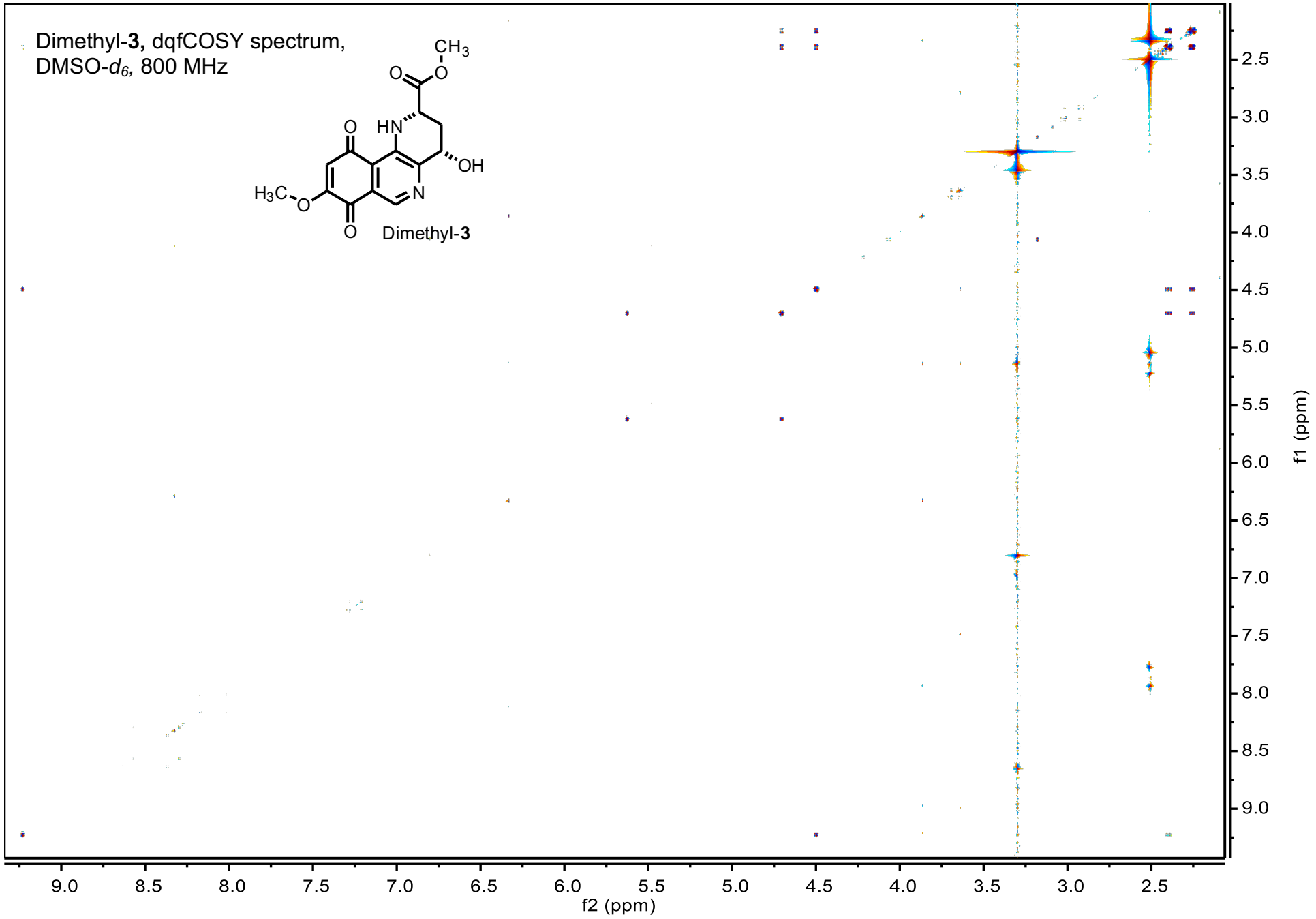


5

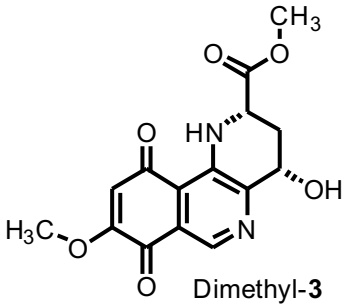
Dimethyl-**3,** ^{1}H NMR spectrum,
DMSO-d_6, 800 MHz



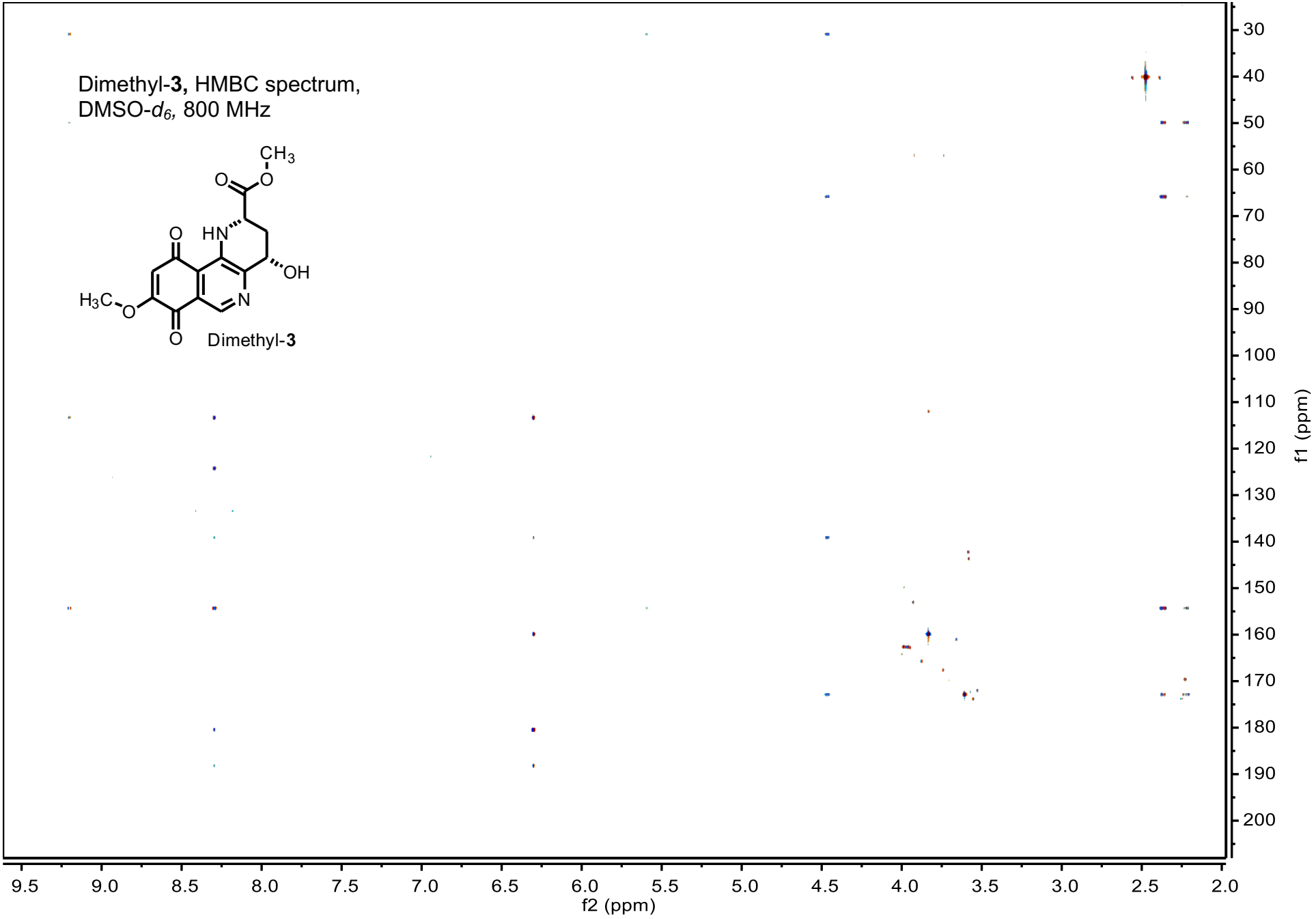
Dimethyl-**3**

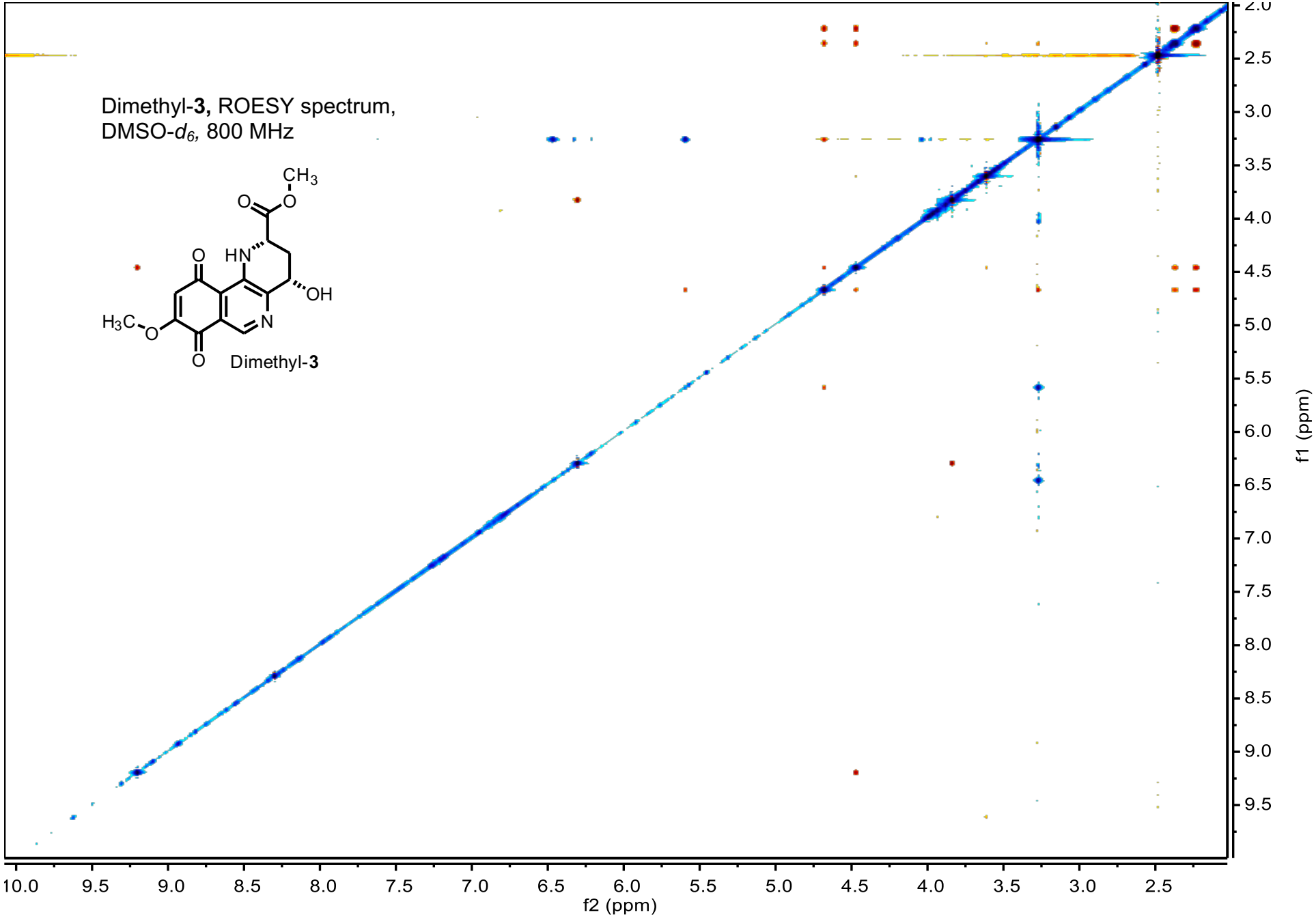

Dimethyl-3, dqfCOSY spectrum,
DMSO-d6, 800 MHz
CH3
O
O
HN
O
OH
H3C
O
N
O
Dimethyl-3
f2 (ppm)
f1 (ppm)
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5

Dimethyl-**3,** coupled HSQC spectrum,
DMSO-d_6, 800 MHz

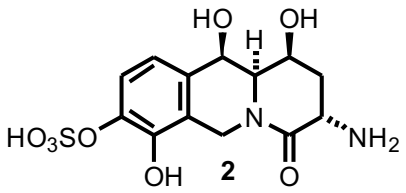


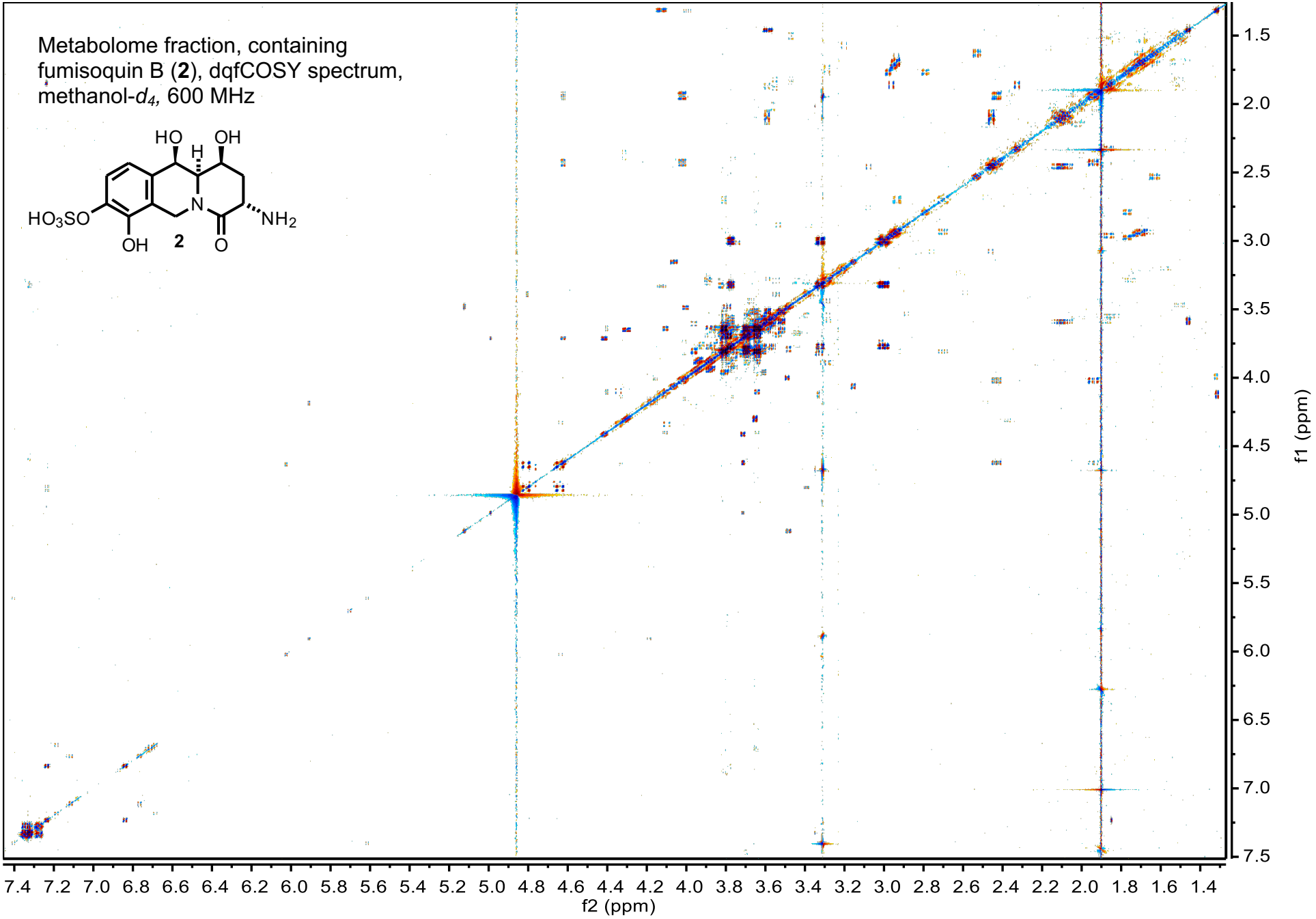
Dimethyl-**3**

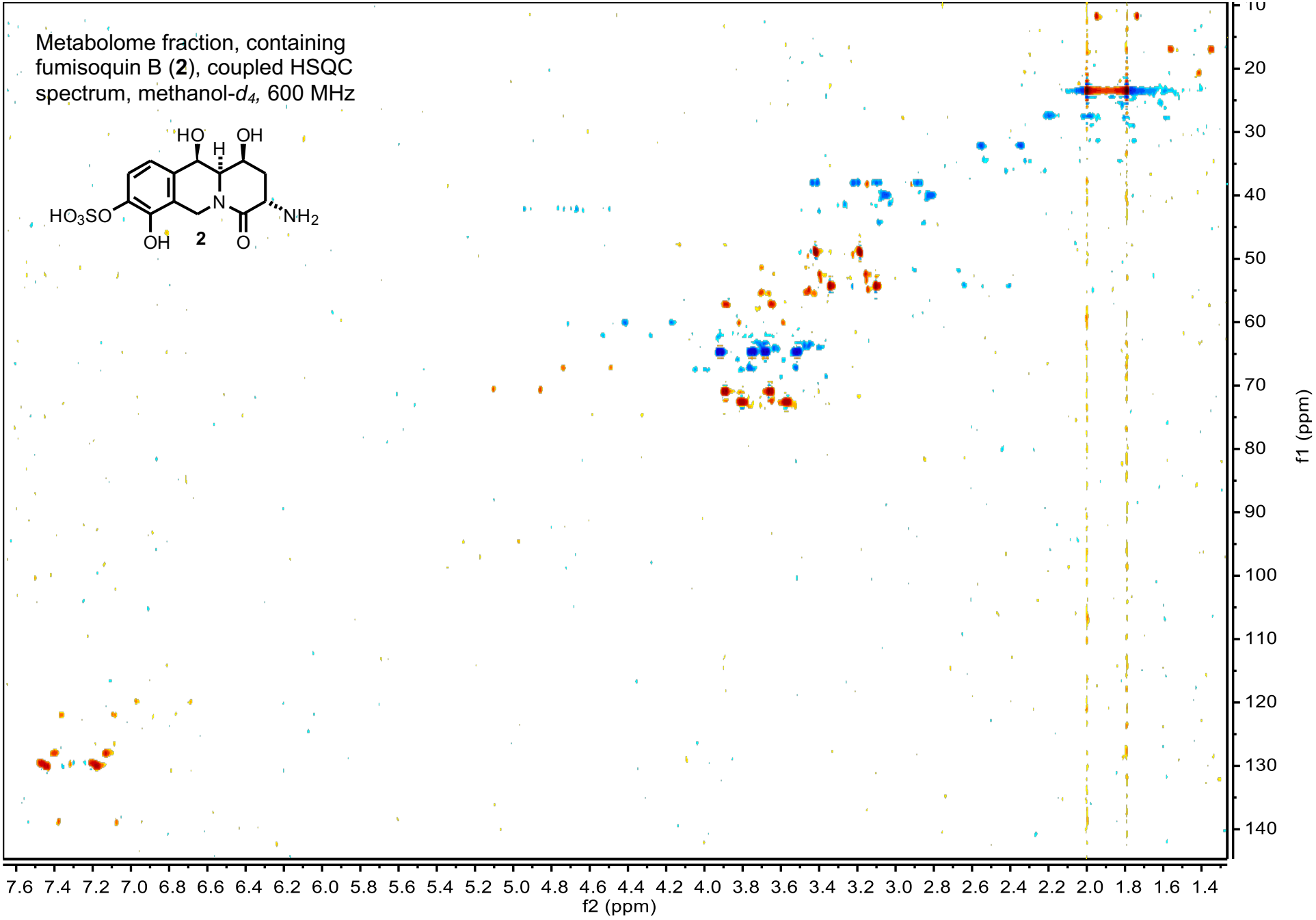

Dimethyl-3, HMBC spectrum,
DMSO-d6, 800 MHz
CH3
O
O
HN
O
OH
H3C
O
N
O
Dimethyl-3
f2 (ppm)
f1 (ppm)

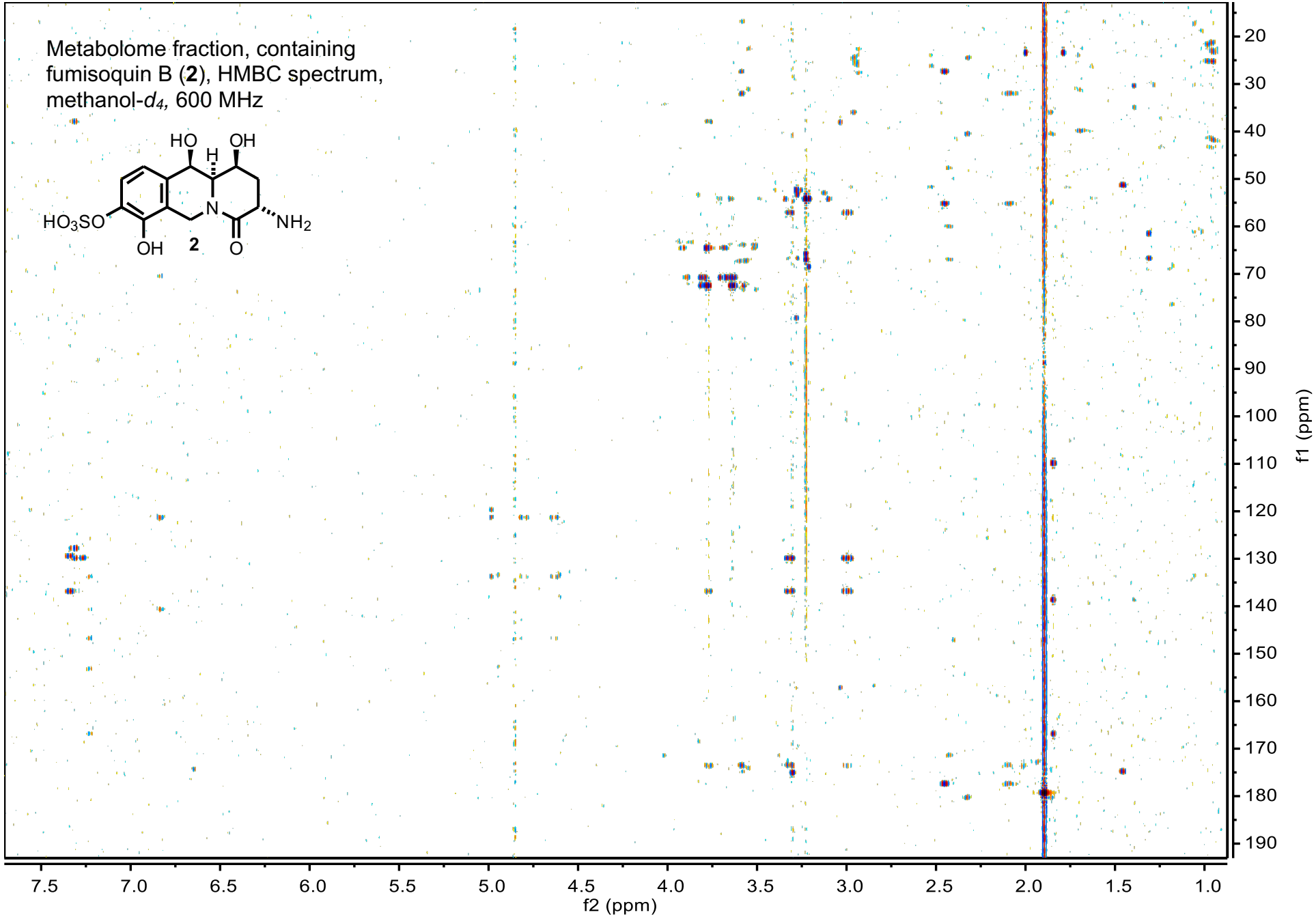

Dimethyl-3, ROESY spectrum,
DMSO-d6, 800 MHz
CH3
O
O
HN
O
OH
H3C
O
N
O
Dimethyl-3
f2 (ppm)
f1 (ppm)

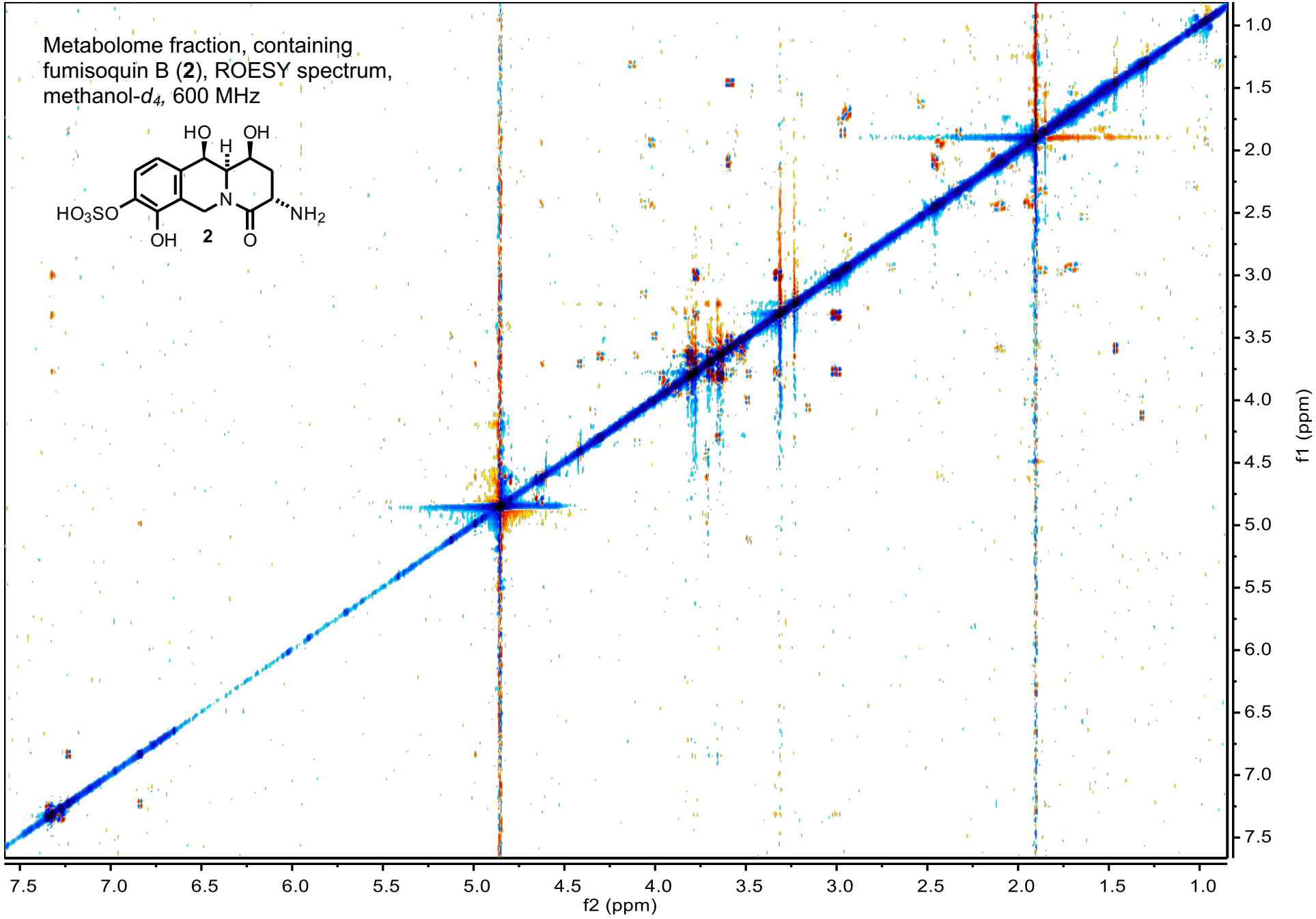
Metabolome fraction, containing fumisoquin B (**2**), ^{1}H NMR spectrum, methanol-d_4, 600 MHz



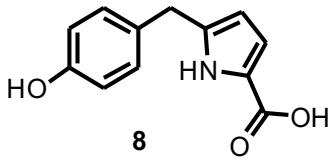

Metabolome fraction, containing fumisoquin B (2), dqfCOSY spectrum, methanol-d4, 600 MHz
HO
H
OH
HO3SO
N
NH2
OH
2
O
f1 (ppm)
f2 (ppm)


Metabolome fraction, containing fumisoquin B (2), coupled HSQC spectrum, methanol-d_4, 600 MHz
HO
H
OH
HO_3SO
N
NH_2
OH
2
O
f2 (ppm)
f1 (ppm)


Metabolome fraction, containing
fumisoquin B (2), HMBC spectrum,
methanol-d4, 600 MHz
HO
OH
H
N
HO3SO
NH2
OH
2
O
f2 (ppm)
f1 (ppm)

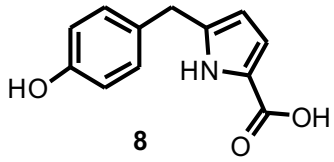

Metabolome fraction, containing
fumisoquin B (2), ROESY spectrum,
methanol-d4, 600 MHz
HO
H
OH
HO3SO
N
NH2
OH
2
O
f2 (ppm)
f1 (ppm)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0

Metabolome fraction containing **8** as major component,
^{1}H NMR spectrum in methanol-d_4, 600 MHz



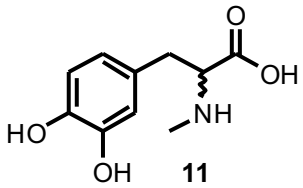
8

Metabolome fraction containing **8** as major component,
HMBC spectrum in methanol-d_4, 600 MHz



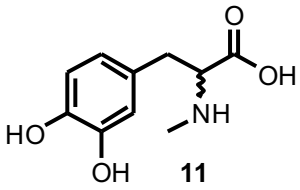
8

Compound **11**, ^{1}H NMR spectrum in D_2O, 600 MHz

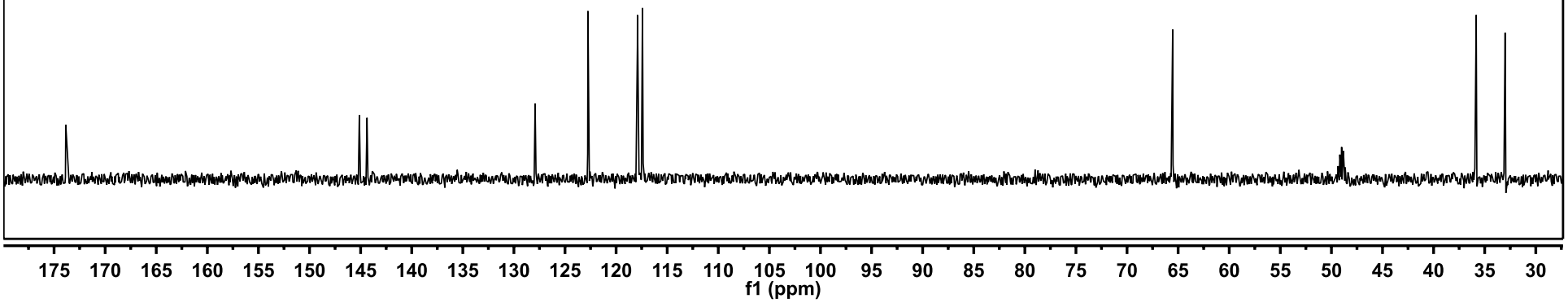


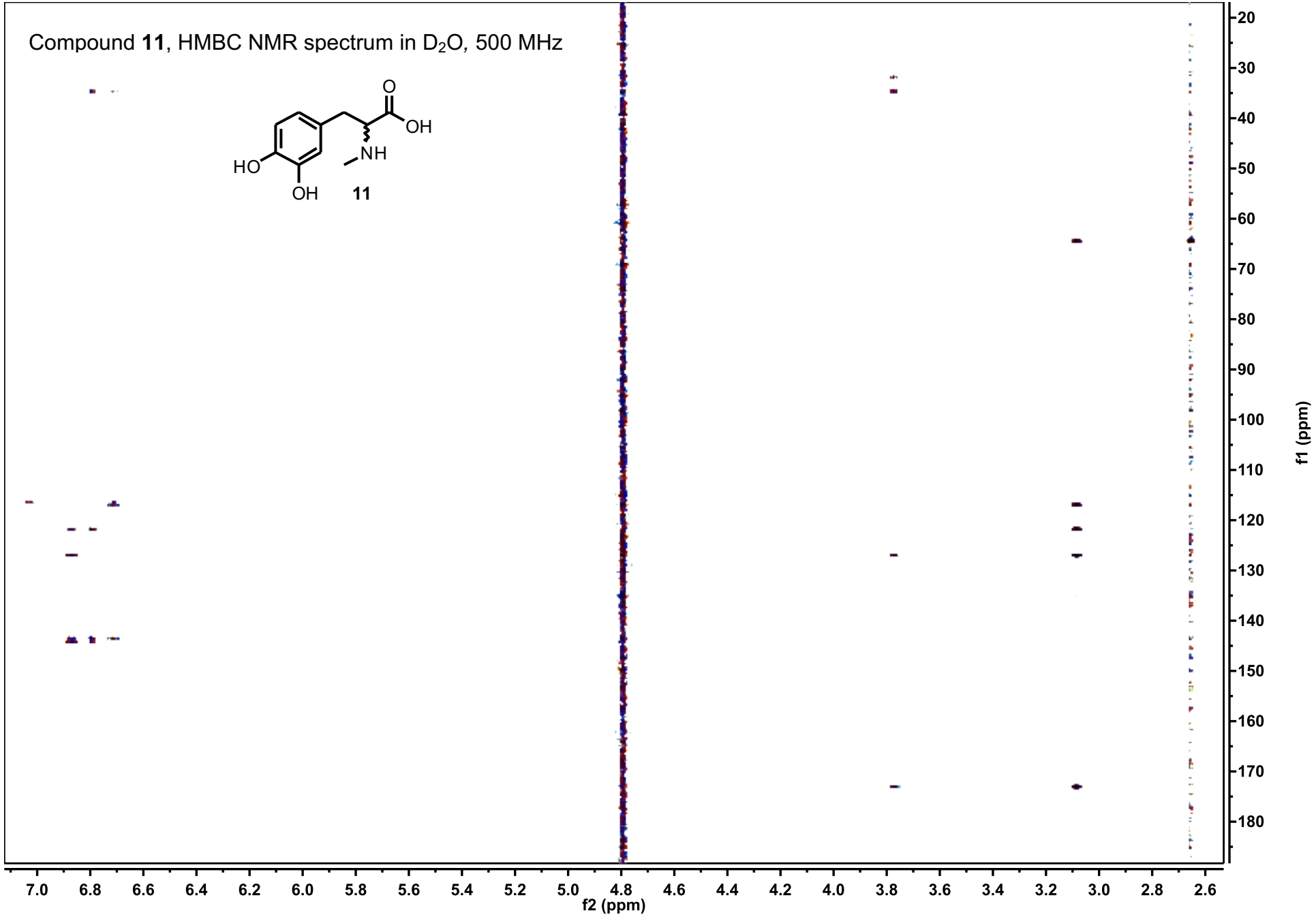
11

Compound **11**, ^{13}C NMR spectrum in D_2O, 500 MHz

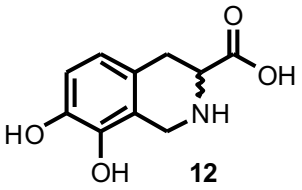


11

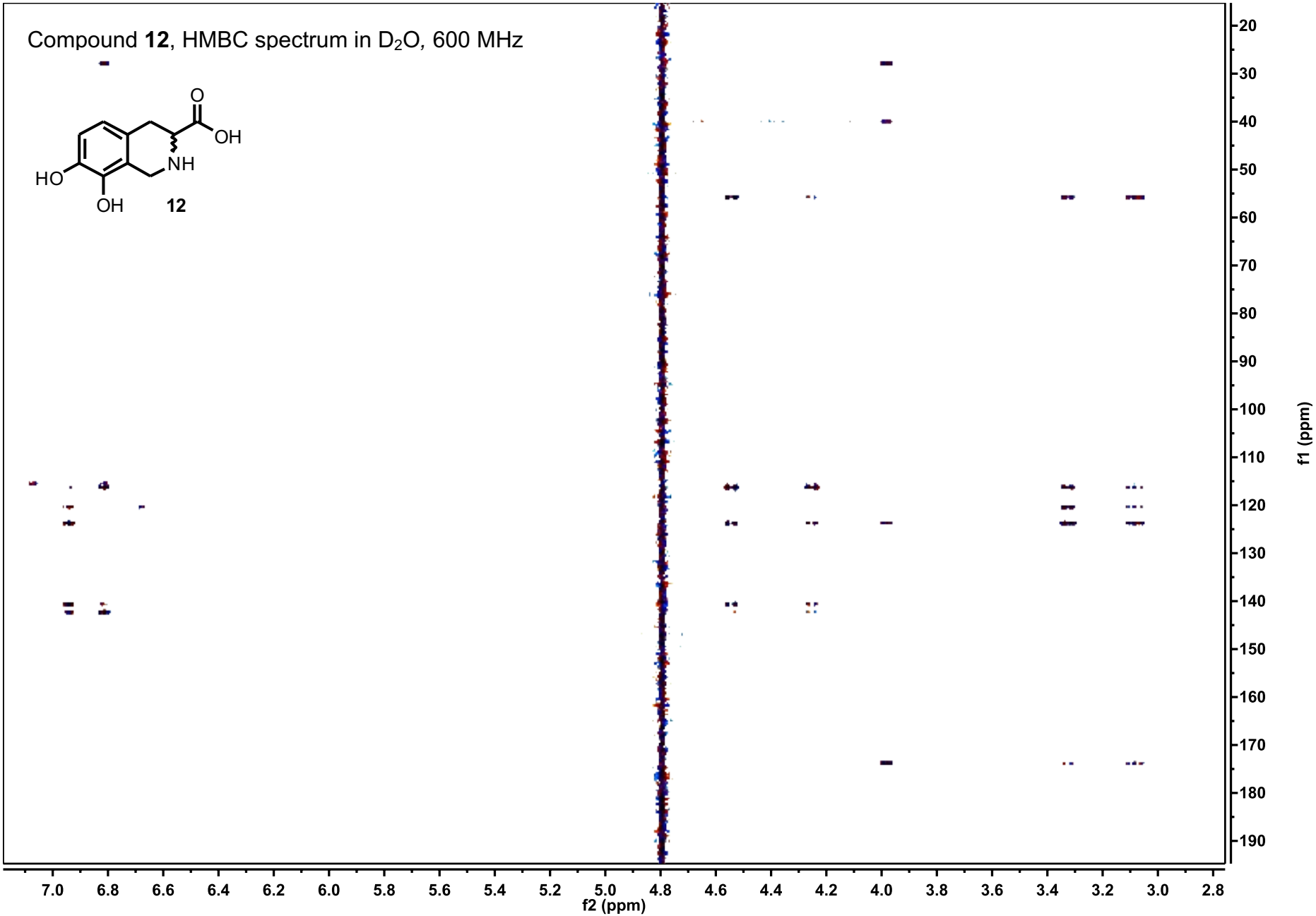



Compound **11**, HMBC NMR spectrum in D_2O, 500 MHz
O
OH
NH
HO
OH
11
f2 (ppm)
f1 (ppm)

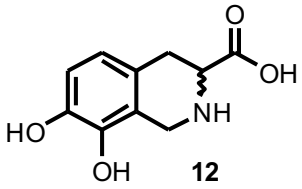
Compound **12**, ^{1}H NMR spectrum in D_2O, 600 MHz



12

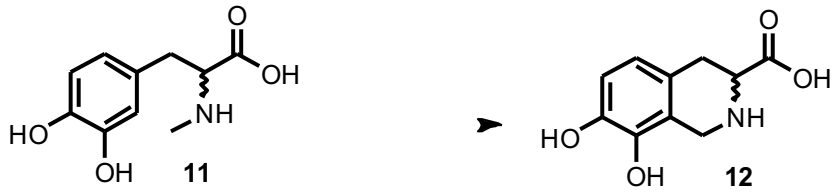

Compound 12, HMBC spectrum in D2O, 600 MHz
O
OH
NH
HO
OH
12
f2 (ppm)
f1 (ppm)

Compound **12**, decoupled HSQC spectrum in D_2O, 600 MHz



12

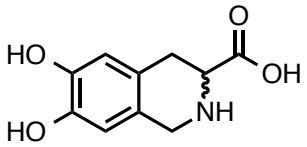
Reaction of **11** with 0.003 equivalents of FsqB in 1 mM phosphate buffer at pH 7, ^{1}H NMR spectrum in D_2O, 600 MHz, also see Figure 5



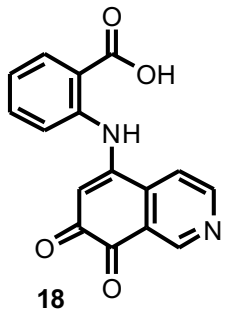
11

12

6,7-Dihydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (**13)** in 1 mM phosphate buffer at pH 7,
^{1}H NMR spectrum water-d_2, 600 MHz

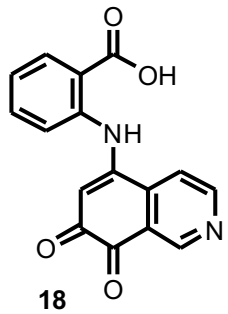


Metabolome fraction, containing shunt metabolite **18** as the major component, ^{1}H NMR spectrum, DMSO-d_6, 600 MHz



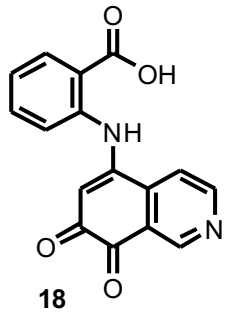
18

Metabolome fraction, containing shunt metabolite **18** as the major component, dqfCOSY spectrum, DMSO-d_6, 600 MHz



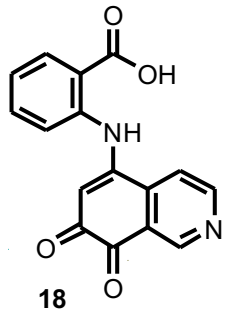
18

Metabolome fraction, containing shunt metabolite **18** as the major component, decoupled HSQC spectrum, DMSO-d_6, 600 MHz

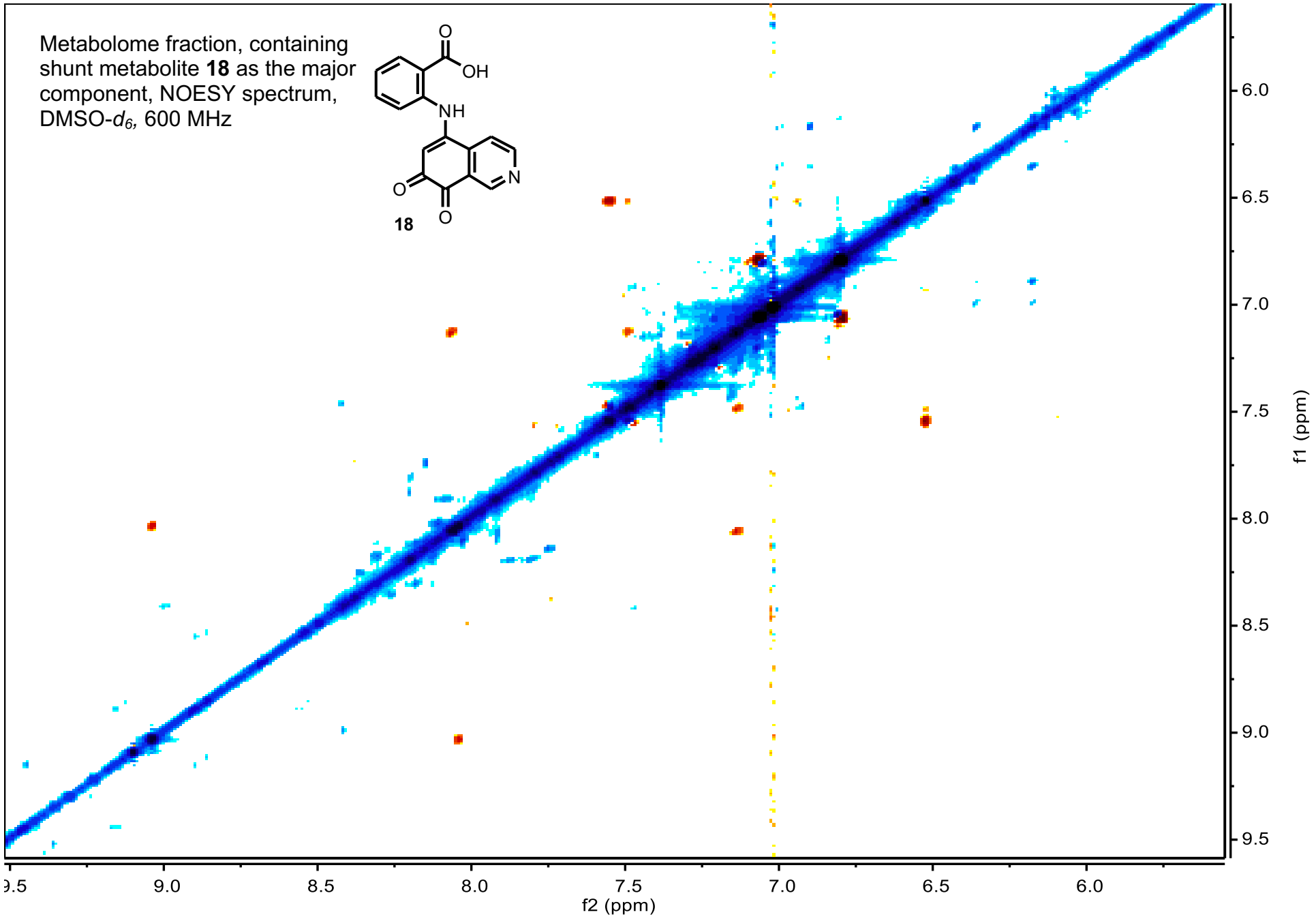


18

Metabolome fraction, containing shunt metabolite **18** as the major component, HMBC spectrum, DMSO-d_6, 600 MHz



18


Metabolome fraction, containing shunt metabolite **18** as the major component, NOESY spectrum, DMSO-d_6, 600 MHz
O
OH
NH
N
O
O
18
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
f1 (ppm)
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
f2 (ppm)

REFERENCES

1. Shimizu, K. & Keller, N. P. Genetic involvement of a cAMP-dependent protein kinase in a G protein signaling pathway regulating morphological and chemical transitions in Aspergillus nidulans. *Genetics* **157**, 591-600 (2001).

2. Sekonyela, R. *et al.* RsmA regulates *Aspergillus fumigatus* gliotoxin cluster metabolites including cyclo(L-Phe-L-Ser), a potential new diagnostic marker for invasive aspergillosis. *PLoS ONE* **8**, e62591, (2013).

3. Palmer, J. M. *et al.* Loss of CclA, required for histone 3 lysine 4 methylation, decreases growth but increases secondary metabolite production in Aspergillus fumigatus. *PeerJ* **1**, e4, (2013).

4. Yu, J. H. *et al.* Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi. *Fungal Genet Biol* **41**, 973-981, (2004).

5. Kalb, D., Lackner, G. & Hoffmeister, D. Functional and phylogenetic divergence of fungal adenylate-forming reductases. *Appl Environ Microbiol* **80**, 6175-6183, (2014).

6. Schneider, P., Bouhired, S. & Hoffmeister, D. Characterization of the atromentin biosynthesis genes and enzymes in the homobasidiomycete Tapinella panuoides. *Fungal Genet Biol* **45**, 1487-1496, (2008).

7. Bok, J. W. *et al.* Chromatin-level regulation of biosynthetic gene clusters. *Nat Chem Biol* **5**, 462-464, (2009).

8. Carpenter, J. F. An improved synthesis of 5,6-diacetoxy-N-methylindole and of epinine. *J Org Chem* **58**, 1607-1609, (1993).

9. Clauser, K. R., Baker, P. & Burlingame, A. L. Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal. Chem.* **71,** 2871–82 (1999).

10. Xue, T., Nguyen, C. K., Romans, A., Kontoyiannis, D. P. & May, G. S. Isogenic auxotrophic mutant strains in the Aspergillus fumigatus genome reference strain AF293. *Arch. Microbiol.* **182,** 346-353 (2004).

11. He, Z., Price, M.S., OBrian, G.R., Georgianna, D.R. & Payne, G.A. Improved protocols for functional analysis in the pathogenic fungus Aspergillus flavus. *BMC Microbiology.* **7**, 104 (2007).

12. Palmer, J. M. *et al.* Loss of CclA, required for histone 3 lysine 4 methylation, decreases growth but increases secondary metabolite production in Aspergillus fumigatus. *PeerJ* **1**, e4 (2013).

13. Sekonyela, R. *et al.* RsmA regulates Aspergillus fumigatus gliotoxin cluster metabolites including cyclo(L-Phe-L-Ser), a potential new diagnostic marker for invasive aspergillosis. *PLoS One* **8**, e62591 (2013).

14. Lim, F.Y., *et al.* Genome-Based Cluster Deletion Reveals an Endocrin Biosynthetic Pathway in Aspergillus fumigatus. *Appl. Environ. Microbiol.* **78**, 4117-4125 (2012).

15. Calvo, A.M., Bok, J.W., Brooks, W. & Keller, N.P. LeA Is Required for Toxin and Sclerotial Production in Aspergillus parasiticus. *Appl. Environ. Microbiol.* **70**, 4733-4739 (2004)

APPENDIX D

BIOSYNTHESIS OF THE STRUCTURALLY UNPRECEDENTED IMIZOQUINS IN THE PLANT PATHOGENIC FUNGUS ASPERGILLUS FLAVUS

Figure D1. Picture of methanol extract from lyophilized OE::*imqK* cultures showing purple pigmentation resulting from imizoquin production.

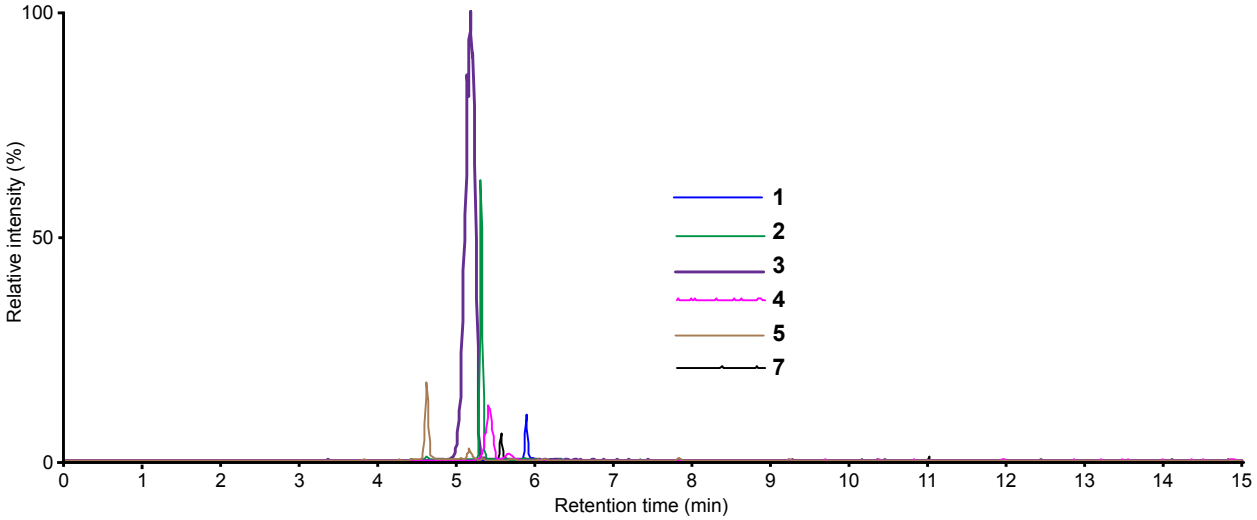


Figure D2. LC-HRMS extracted-ion-chromatograms for *imq*-associate metabolites.

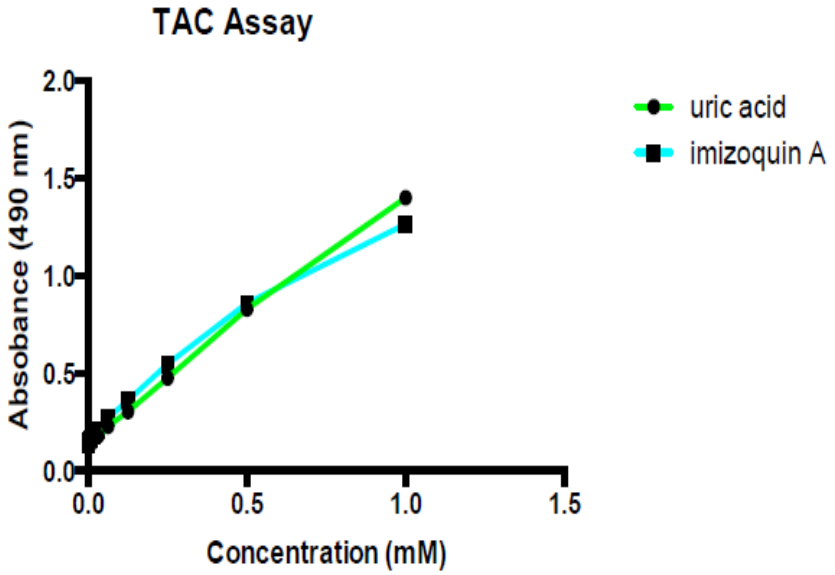


Figure D3. Total antioxidant capacity assay (TAC) with uric acid and imizoquin A (**1**).

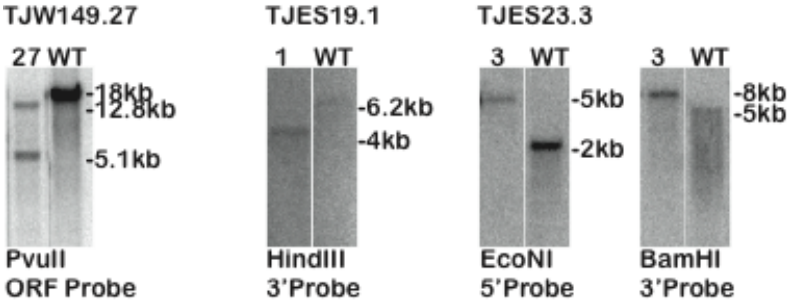


Figure D4. Southern analysis confirmation of all mutants used in this study. Mutant (number) lanes are on the left of each image and the parental strain ("WT") is shown on the right. Expected band sizes correlating with those seen in the images are marked accordingly. Under each image is the enzymes used in the restriction digest as well as the nucleic acid probe used.

Strains, media, and growth conditions. The fungal strains used in this study are listed in **Table D2**. All strains were grown at 37 °C on glucose minimal medium (GMM)[1] and, when appropriate, were supplemented with 0.56 g/L uracil, 1.26 g/L uridine, 1.0 g/L arginine and maintained as glycerol stocks at -80 ºC. *Escherichia coli* strain DH5α was propagated in LB medium with appropriate antibiotics for plasmid DNA.

Gene cloning, plasmid construction, and genetic manipulations. The oligonucleotide sequences for PCR primers are given in **Table D3**. PCR amplification was carried out on a C1000^{TM} Thermal Cycler (Bio-Rad).

Construction of the *A. flavus* Δ*ku70* strain was generated using standard double joint PCR procedures[4]. Briefly, *A. parasiticus pyrG* (*A.ppyrG)* was amplified from pJW24. Then, an approximately 1 kb fragment upstream and downstream of the *ku70* gene was amplified from genomic DNA of *A. flavus* NRRL3357 using designated primers, respectively. These three amplified PCR products were purified with a QIAquick gel extraction kit (Qiagen, Valencia, CA), quantified, and fused using standard double joint PCR procedures. The final PCR product was amplified with the bottom primer pairs, Fku70IF and Fku70IR, confirmed with endonuclease digestion, and purified for fungal transformation. The first two rounds of PCR were done with Pfu Ultra II DNA Polymerases (Agilent, Santa Clara, CA), and the final PCR step used Expand long template PCR system (Roche, Indianapolis, IN) according to the manufacturer's instructions. NRRL3357.5 (auxotroph *A. flavus, pyrG*$^{-}$) was used to make the Δ*ku70* mutant. Transformants were screened with PCR then confirmed with Southern digests (Fig S7).

Recycling of the *pyrG* auxotrophy in the TJW149.27 background was done by generating a marker-less cassette to remove the *A.ppyrg* marker from the ∆ku70 locus using standard single single joint PCR procedures. Briefly approximately 1 kb flanks upstream and downstream of the inserted *A.ppyrG* were amplified from genomic DNA of TJW149.27 using designated primers and subsequently purified with a QIAquick gel extraction kit. These fragments were fused using standard single-joint reaction procedures and amplified with the bottom primer pairs AFLA_pyrGrecyc_5'F and AFLA_pyrGrecyc_3'R, confirmed with endonuclease digestion prior to transformation. Five µg of the double-joint PCR cassette was used to delete *A.ppyrG* by using TJW149.27 as the recipient host. After transformation, transformants were grown on minimal media plates with 1mg/ml 5-fluoroorotic acid and 10mM uracil and uridine supplements for selection of pyrG auxotrophic strains. All strains were verified by PCR and Southern blot analysis (**Figure D4**). Multiple confirmed strains for each mutant were stored at -80 °C with 33 % glycerol for future use.

Construction of *A. flavus* OE *imqA* cassette was generated using standard double joint PCR procedures[4]. Briefly, *A. parasiticus pyrG* (*A.ppyrG)* and *A. nidulans* (p)gpdA was amplified from pJMP9.1 Then, an approximately 1 kb fragment upstream and downstream of the *imqA* gene was amplified from genomic DNA of *A. flavus* NRRL3357 using designated primers. These three amplified PCR products were purified with a QIAquick gel extraction kit (Qiagen, Valencia, CA), quantified, and fused using standard double joint PCR procedures. The final PCR product was amplified with the bottom primer pairs, AFLA_OE064330_5'F and

AFLA_OE064330_3'R, confirmed with endonuclease digestion, and purified for fungal transformation. The first two rounds of PCR were done with Pfu Ultra II DNA Polymerases (Agilent, Santa Clara, CA), and the final PCR step used Expand long template PCR system (Roche, Indianapolis, IN) according to the manufacturer's instructions. TJE19.1 (auxotroph *A. flavus, pyrG*$^{-}$*,* Δ*ku70*) was used to make the *OE*::*imqA*. Deletion cassette of *imqA* was constructed by using double joint PCR with *A. parasiticus pyrG* (*A.ppyrG*) for replacement of *imqA*. The *A.ppyrG* gene was amplified from the plasmid pJMP9.1 as template. Five µg of the double-joint PCR cassette was used to delete *imqA* by again using TJES19.1 as the recipient host. After transformation, transformants were grown on minimal media plates without supplements for screening. All strains were verified by PCR and Southern blot analysis (**Figure D4**). Multiple confirmed strains for each mutant were stored at -80 °C with 33 % glycerol for future use.

Nucleic acid analysis. Preparation of plasmids, restriction enzyme digestions, gel electrophoresis, blotting, probe preparation and hybridization were carried out by standard protocols. *Aspergillus* DNA was extracted using a previously described method[7]. Sequence data were analyzed using the LASERGENE software package from DNASTAR.

Northern analysis. 50 mL of liquid GMM[1] were inoculated with 1.0×10^{6} spores (asexual) mL^{-1} of all appropriate strains in this study and incubated with shaking at 250 rpm at 25 °C. After 48 h, the mycelium was collected and total RNA was extracted by using Isol-RNA Lysis Reagent according to the manufacturer's instructions (5 Prime).

Fermentation and metabolome extraction for comparative metabolomics by LC-MS.

Preparation for HPLC-MS analysis: *A*. *fumigatus* strains were inoculated (1.0×10^6 spores mL^{-1}) into 50 mL GMM[1] in a 125 mL Erlenmeyer flask at 37 °C with shaking at 220 rpm. After 4 days, liquid fungal cultures including fungal tissue and media were frozen using a dry ice acetone bath, and lyophilized. The lyophilized residues were extracted with 30 mL of MeOH for 1.5 h with vigorous stirring. Extracts were filtered over cotton, evaporated to dryness, and stored in 4 mL vials. Crude extracts were suspended in 0.5 mL of MeOH and centrifuged to remove insoluble materials, and the supernatant was subjected to LC-MS analysis.

Analytical methods and equipment overview. (a) NMR spectroscopy: NMR spectroscopic instrumentation: a Bruker Avance^{III} HD (800 MHz ^{1}H reference frequency, 201 MHz for ^{13}C) equipped with a 5 mm CPTCL ^{1}H-^{13}C/^{15}N cryo probe. Non-gradient phase-cycled dqfCOSY spectra were acquired using the following parameters: 0.6 s acquisition time, 400-600 complex increments, 8, 16 or 32 scans per increment. Non-gradient HSQC, HMQC, and HMBC spectra were acquired with these parameters: 0.25 s acquisition time, 200-500 increments, 8-64 scans per increment. ^{1}H, ^{13}C-HMBC spectra were optimized for $J_{H,C}$ = 6 Hz. HSQC spectra were acquired with or without decoupling. NMR spectra were processed and baseline corrected using MestreLabs MNOVA software packages. (b) Mass spectrometry: LC-HRMS was performed on a Thermo Scientific-Dionex Ultimate3000 UHPLC system equipped with a diode array detector and connected to a Thermo Scientific Q Exactive Orbitrap operated in electrospray positive (ESI^+)

or electrospray negative (ESI^-) ionization mode. Low-resolution LC-MS was performed on an Agilent 1100 series HPLC system equipped with a diode array detector and connected to a Quattro II mass spectrometer (Micromass/Waters) operated in ESI^+ or ESI^- mode. Data acquisition and processing for the LC-HRMS was controlled by Thermo Scientific Xcalibur software. Data acquisition and processing for the LC-MS was controlled by Waters MassLynx software. (c) Chromatography: flash chromatography was performed using a Teledyne ISCO CombiFlash system. For semi-preparative HPLC Agilent Zorbax Eclipse XDB-C18 or -C8 columns (25 cm x 10 mm, 5 µm particle diameter) were used. An Agilent Zorbax RRHD Eclipse XDB-C18 column (2.1 x 100 mm, 1.8 µm particle diameter) was used in the LC-HRMS *A. flavus* mutant profiling analysis. An Agilent Zorbax Eclipse XDB-C18 column (4.6 x 250 mm, 5 µm particle diameter) was used in the LC-MS *A. flavus* mutant profiling analysis.

Chromatographic enrichment of compounds 1-7. Methanol extracts derived from 1-4 L of *A. flavus* cultures were fractionated using large-scale reverse-phase flash chromatography on a Teledyne ISCO CombiFlash with a Teledyne C18 gold (100 gram) column with acetonitrile (organic phase) and 0.1 % acetic acid in water (aqueous phase) as solvents at a flow rate of 60 mL/min. A linear ramp from 0 % organic to 100 % organic over 30 min was used and fractions containing compounds of interest were collected, evaporated to dryness, and stored in 8 mL glass vials at -20 °C. Fractions containing compounds **1-7** were further purified via semi-preparative HPLC using an Agilent XDB C-18 or C-8 column (25 cm x 10 mm, 5 µm particle diameter) acetonitrile (organic phase) and 0.1 % acetic acid in

water (aqueous phase) as solvents at a flow rate of 3.6 mL/min. A solvent gradient scheme was used, starting at 5 % organic for 3 min, followed by a linear increase to 100 % organic over 27 min, holding at 100 % organic for 5 min, then decreasing back to 5 % organic for 0.1 min, and holding at 5 % organic for the final 4.9 min, for a total of 40 min.

Table D1. LC-HRMS data of reported compounds

Compound	HR-ESI(+/-) Observed (m/z)	Ion	Calculated Ion Formula	Calculated m/z	Retention time [min]
1	565.1942	$[M-H]^-$	$C_{28}H_{29}N_4O_9^-$	565.1940	5.79
2	581.1889	$[M-H]^-$	$C_{28}H_{29}N_4O_{10}^-$	581.1889	5.25
3	569.2253	$[M-H]^-$	$C_{28}H_{33}N_4O_9^-$	569.2253	4.97
4	553.2304	$[M-H]^-$	$C_{28}H_{33}N_4O_8^-$	553.2304	5.22
6	553.2048	$[M-H]^-$	$C_{28}H_{31}N_4O_{10}^-$	553.2046	4.48
7	420.1203	$[M-H]^-$	$C_{22}H_{18}N_3O_6^-$	420.1201	5.50

Table D2. Fungal strains and plasmids used in this study

***Aspergillus flavus* NRRL3357 background strains**		
NRRL3357.5	*pyrG1*	2
TJW149.27	*Δku70::pyrG*	This study
TJES19.1	*ΔpyrG, Δku70*	This study
TJES 23.3	*A.ppyrG::gpdA(p)::imqA, Δku70*	This study
TJES 27.1	*ΔimqA::A.ppyrG, Δku70*	This study

Table D3. PCR primer sets used in this study

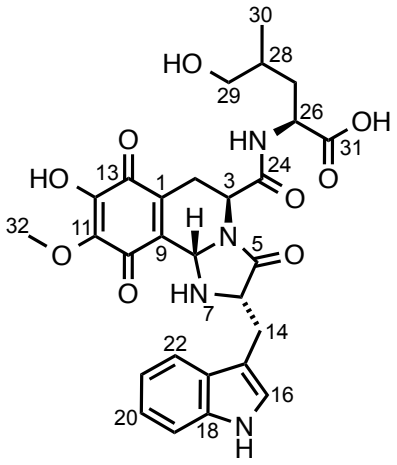
Name of the primer	Oligonucleotide sequence (5'-3')	Uses
flvku70F5	ACATCTCTTCCGTCAAAGGCGC	Amplification of *ku70* 5' flanking region
flavku70R5	CGATATCAAGCTATCGATACCTCGACT CTGTGTTGAGAGTCGTAAGTCATGAAT TGCG	Amplification of *ku70* 5' flanking region
flvku70F3	GTCGCTGCAGCCTCTCCGATTGTCGAA TGACAACGCTAGTATTGGTTACGAGAG ACAG	Amplification of *ku70* 3' flanking region
flavku70R3	AGAATGGCTACGTCAACCTCCG	Amplification of *ku70* 3' flanking region
Fku70IF	ATGAGGAAGAGGAGGAGACCG	

Fku70IR	CACTTTTCAATCGTGCGAGCCG	Amplification of Δ*ku70* cassette
AFLA_pyrGrecyc 3'R	CAATCGCTTTCCTCCATGCC	Amplification of Δ*ku70* cassette
AFLA_ pyrGrecyc 3'F	CGTGTGGAGGAGTATTTGGAGCAGAA ATGACCTAGGATGATGCTGTAGGG	Amplification of Δ*ku70* 3'flank, pyrG recycling
AFLA_ pyrGrecyc 5'R	TCATTTCTGCTCCAAATACTCCTCCACA CGGCAAAAGTGCCGTTGAACGG	Amplification of Δ*ku70* 3'flank, pyrG recycling
AFLA_ pyrGrecyc 5'F	ACTGCCTGTAGTCGAACAGG	Amplification of Δ*ku70* 5'flank, pyrG recycling
AFLA_OE064330_ 5'R	GCACACCCTCGGAATAGTCCTCTCGGG CCATTCCTTGACCCAGGTGTGGC	Amplification of Δ*ku70* 5'flank, pyrG recycling
AFLA_OE064330_5'F	AGAAACCACCAGCCTCAGCG	Amplification of *imqJ* 5'flank, OE
AFLA_OE064330_3'R	TCGGAGTTGGAGCTTCCTCG	Amplification of *imqJ* 5'flank, OE
AFLA_OE064330 _3'F	TGGGGCGCGCCGAGGGCTATGTGCAG GAGATAAGCCCCCGCATTGCATCG	Amplification of *imqJ* 3'flank, OE
AFLA_KO064330_5'F	AAGGTTCCATCGCGGTCAGC	Amplification of *imqJ* 3'flank, OE
AFLA_KO064330_5'R	AGCGCCAATTGCTGTTGCCAGGTGAG GGTATTTGGCGAGATCAGCGACGG	Amplification of *imqJ* 5'flank, OE
AFLA_KO064330_3'F	CTTGGCATCACGCATCAGTGCCTCCTC TCAAAGTCCCTCAGATGCGGTGC	Amplification of *imqJ* 5'flank, OE
AFLA_KO064330_5'R	TGACACGCACACAACCCTGG	Amplification of *imqJ* 3'flank, OE
AFLA064220_FOR_Nprobe	TAGCTGAGCTCTTCCGACGG	Amplification of *imqJ* 3'flank, OE
AFLA064220_REV_Nprobe	TCCTGGTGGTCTCTAGTGCC	Northern probe
AFLA064230_FOR_Nprobe	ATTCCGTATTCCCTGCACCC	Northern probe
AFLA064230_REV_Nprobe	GCTCGGCATCAATAAGGTCG	Northern probe
AFLA064240_FOR_Nprobe	AGACCAACTGGCCTACCAGC	Northern probe
AFLA064240_REV_Nprobe	TCACTTGCCGCGTGATGACC	Northern probe
AFLA064250_FOR_Nprobe	GTGGAGTGGAGAAAGCAGCC	Northern probe
AFLA064250_REV_Nprobe	TACTCACGAACCGAGTCGCC	Northern probe
AFLA064260_FOR_Nprobe	TCTCTTGTGCCATCCGTGGC	Northern probe
AFLA064260_REV_Nprobe	TTATCCTCCCGATCGATCGC	Northern probe
AFLA064270_FOR_Nprobe	AAGGAGACACTGACCAGCGG	Northern probe
AFLA064270_REV_Nprobe	ATACAACGCGGGTCTCTCCG	Northern probe
AFLA064280_FOR_Nprobe	TCTTTTCCCCAGCGTGGACG	Northern probe
AFLA064280_REV_Nprobe	TGAGGCACAAGGACCTGTCG	Northern probe
AFLA064290_FOR_Nprobe	AGGACATCTCCGCCTTTCCC	Northern probe
AFLA064290_REV_Nprobe	ACATTCTCGACCGTCCTGCG	Northern probe
AFLA064300_FOR_Nprobe	TCTTCCCATAGGCCGGATGG	Northern probe

AFLA064300_REV_Nprobe	TTCCCGCATCATTCGCAGCG	Northern probe
AFLA064310_FOR_Nprobe	AGGGTTAGGGATGGTCACGG	Northern probe
AFLA064310_REV_Nprobe	ACTGTTCTGCAACCACGGCC	Northern probe
AFLA064320_FOR_Nprobe	ACCAGGCGTTCATGCTGTGG	Northern probe
AFLA064320_REV_Nprobe	ATGCCACACACCAGGAACCC	Northern probe
AFLA064330_FOR_Nprobe	TAAGCCCCCGCATTGCATCG	Northern probe
AFLA064330_REV_Nprobe	TGCTTGAATGGTCCGGACCC	Northern probe
AFLA064340_FOR_Nprobe	GAATGGAGAAGCTGCATCGG	Northern probe
AFLA064340_REV_Nprobe	ACTTCGATCAGACCCGAGCC	Northern probe
AFLA064350_FOR_Nprobe	ATGCGCCGATTGACGAGTGG	Northern probe
AFLA064350_REV_Nprobe	TGGATGAGACGCCATTGCGC	Northern probe
		Northern probe

Table D4. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for compound **1** in methanol-d_4.

Chemical shifts were referenced to δ(C<u>H</u>D$_2$OD) = 3.31 ppm and δ(^{13}<u>C</u>HD$_2$OD$_3$) = 49.00.^{13}C chemical shifts were determined via HMBC, HSQC and direct observation ^{13}C spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. NOESY correlations were observed using a mixing time of 650 ms. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

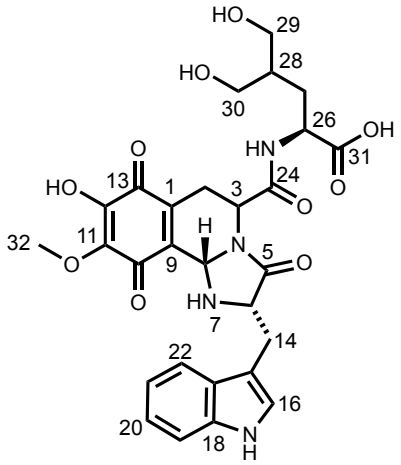


No.	δ_c	Proton	δH(J_{HH}[Hz])	HMBC	NOESY
1	134.65				
2	23.62	2-H$_a$	2.67($J_{2a,2b}$ = 17.9) ($J_{2a,3}$ = 1.5) ($J_{2a,8}$ = 5.2)	1,3,9,13,23	
		2-H$_b$	1.54($J_{2b,2a}$ = 17.9) ($J_{2b,3}$ = 7.5) ($J_{2b,8}$ = 5.1)	1,3,9,10w,23	8
3	50.64	3-H	4.90($J_{3,2a}$ = 1.5) ($J_{3,2b}$ = 7.5)	1,2,5,8,23	
4					
5	176.44				
6	61.21	6-H	4.20($J_{6,14a}$ = 6.0) ($J_{6,14b}$ = 4.1) ($J_{6,8}$ = 5.2)	5,8,14,15	8
7		7-NH			
8	66.84	8-H	5.53($J_{8,2a}$ = 5.2) ($J_{8,2b}$ = 5.1) ($J_{8,6}$ = 5.2)	1	2b,6
9	138.24				
10	183.50				
11	139.23				
12	145.92				
13	184.22				
14	30.01	14-H$_a$	3.05($J_{14a,14b}$ = 13.8) ($J_{14a,6}$ = 6.0)	5,6,15,16	19,21
		14-H$_b$	3.12($J_{14b,14a}$ = 13.8) ($J_{14b,6}$ = 4.1)	5,6,15,16	
15	111.06				

16	125.27	16-H	6.93	14,15,18,23	14a/b
17		17-NH			
18	137.66				
19	112.41	19-H	7.12($J_{19,20}$ = 7.6)	21,23	
20	122.37	20-H	6.99($J_{20,19}$ = 7.6) ($J_{20,21}$ = 5.8)	18,22	
21	119.99	21-H	6.89($J_{21,20}$ = 5.8) ($J_{21,22}$ = 7.8)	19,23	
22	120.53	22-H	7.50($J_{22,21}$ = 7.8)	18,20	14a/b
23	129.47				
24	171.24				
25		25-NH			
26	52.23	26-H	4.28($J_{26,27a}$ = 5.2) ($J_{26,27b}$ = 10.5)	24,27,28,31	27a/b,28,29, 30
27	35.89	27-H_a	1.57($J_{27a,27b}$ = 17.3) ($J_{27a,26}$ = 5.2) ($J_{27a,28}$ = 10.5)	26,28,29,30	26,28,29,30
		27-H_b	1.77($J_{27b,27a}$ = 17.3) ($J_{27b,26}$ = 5.2) ($J_{27b,28}$ = 10.5)	26,28,29,30,31	
28	33.87	28-H	1.57($J_{28,29}$ = 6.1) ($J_{28,30}$ = 6.7)	27,29,30	26,27a/b,29, 30
29	68.18	29-H_2	3.31($J_{29,28}$ = 6.1)	27,28,30	26,27a/b,28, 30
30	16.02	30-H_3	0.84($J_{30,28}$ = 6.7)	27,28,29	26,27a/b,28, 29
31	175.92				
32	60.71	32-H_3	3.86	11	

Table D5. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for compound **2** in methanol-d_4.

Chemical shifts were referenced to δ($C\underline{H}D_2OD$) = 3.31 ppm and δ($^{13}\underline{C}HD_2OD_3$) = 49.00.$^{13}C$ chemical shifts were determined via HMBC, HSQC and direct observation ^{13}C spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

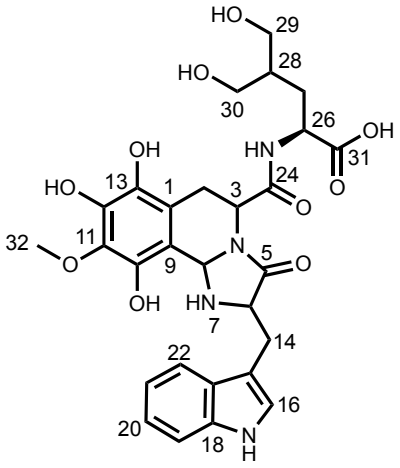


No.	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1	134.20			
2	23.08	2-H_a	2.70 ($J_{2a,2b}$ = 17.9)	1,3,9,13,23
		2-H_b	1.56 ($J_{2b,2a}$ = 17.9) ($J_{2b,3}$ = 7.6)	1,3,9,10w,23
3	50.41	3-H	4.91 ($J_{3,2b}$ = 7.6)	1,2,5,8,23
4				
5	175.98			
6	60.86	6-H	4.22 ($J_{6,14a}$ = 4.4) ($J_{6,14b}$ = 4.4)	5,8,14,15
7		7-NH		
8	66.43	8-H	5.54	1
9	138.02			
10	184.24			
11	147.69			
12	147.14*			
13	182.83			
14	29.67	14-H_a	3.08 ($J_{14a,14b}$ = 14.7) ($J_{14a,6}$ = 4.4)	5,6,15,16
		14-H_b	3.13 ($J_{14b,14a}$ = 14.7) ($J_{14b,6}$ = 4.4)	5,6,15,16
15	110.69			

16	124.94	16-H	6.94	14,15,18,23
17		17-NH		
18	137.34			
19	112.09	19-H	7.14 ($J_{19,20}$ = 8.2)	21,23
20	122.10	20-H	7.03 ($J_{20,19}$ = 8.2) ($J_{20,21}$ = 5.8)	18,22
21	119.71	21-H	6.91 ($J_{21,20}$ = 7.5) ($J_{21,22}$ = 7.5)	19,23
22	120.20	22-H	7.51 ($J_{22,21}$ = 7.5)	15,18,20
23	134.20			
24	170.43			
25		25-NH		
26	52.92	26-H	4.29 ($J_{26,27a}$ = 4.5) ($J_{26,27b}$ = 9.7)	24,27,28,31
27	31.40	27-H_a	1.67 ($J_{27a,27b}$ = 16.5) ($J_{27a,26}$ = 4.5) ($J_{27a,28}$ = 9.7)	26,28,29,30
		27-H_b	1.82 ($J_{27b,27a}$ = 16.5) ($J_{27b,26}$ = 4.5) ($J_{27b,28}$ = 9.7)	26,28,29,30,31
28	41.12	28-H	1.58	26,27,29,30
29	63.76	29-H_a	3.51	27,28,30
		29-H_b	3.55	27,28,30
30	62.32	30-H_a	3.51	27,28,29
		30-H_b	3.55	27,28,29
31	176.35			
32	56.15	32-H_3	3.84	11

Table D6. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for compound **5** in methanol-d_4.

Chemical shifts were referenced to δ(C<u>H</u>D_2OD) = 3.31 ppm and δ(^{13}<u>C</u>HD_2OD_3) = 49.00.^{13}C chemical shifts were determined via HMBC, HSQC and direct observation ^{13}C spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

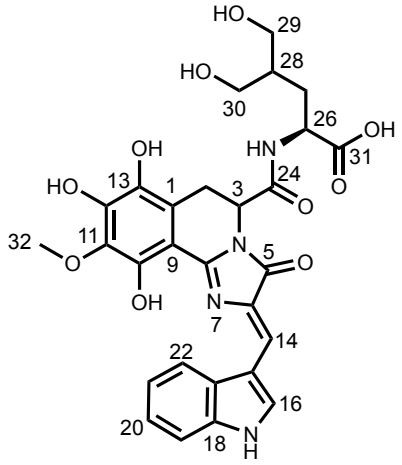


No.	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1	116.22			
2	24.65	2-H_a	3.04	1,3,9,12,13,23
		2-H_b	3.14	1,3,9,12,13,23
3	53.06	3-H	4.64	1,2,5,8,23
4				
5	175.81			
6	60.72	6-H	4.24	5,8,14,15
7		7-NH		
8	68.95	8-H	5.84	1,5,6,9,10
9	112.45			
10	142.19			
11	138.96			
12	140.23			
13	136.39			
14	28.73	14-H_a	2.81	5
		14-H_b	3.30	5
15	111.20			

16	124.30	16-H	7.19	14,15,18,23
17		17-NH		
18	137.87			
19	111.84	19-H	7.34	21,23
20	122.09	20-H	7.10	18,22
21	119.37	21-H	7.01	19,23
22	118.90	22-H	7.58	15,18,20
23	128.07			
24	171.72			
25		25-NH		
26	52.88	26-H	4.44	24,27,28,31
27	31.48	27-H_a	1.71	26,28,29,30
		27-H_b	1.84	26,28,29,30,31
28	40.92	28-H	1.61	26,27,29,30
29	62.32	29-H_a	3.53	27,28,30
		29-H_b	3.60	27,28,30
30	63.73	30-H_a	3.53	27,28,29
		30-H_b	3.60	27,28,29
31	176.54			
32	60.33	32-H_3	3.87	11

Table D7. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for compound **6** in methanol-d_4.

Chemical shifts were referenced to δ(C<u>H</u>D_2OD) = 3.31 ppm and δ(^{13}<u>C</u>HD_2OD_3) = 49.00.^{13}C chemical shifts were determined via HMBC, HSQC and direct observation ^{13}C spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

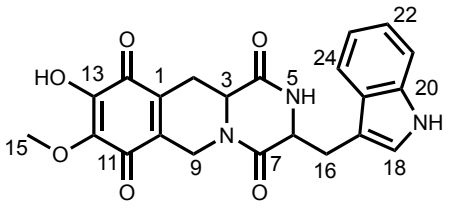


No.	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1	115.87			
2	25.63	2-H_a	3.22 ($J_{2a,2b}$ = 16.7) ($J_{2a,3}$ = 8.0)	1,3,8w,9,10,12,13,24
		2-H_b	3.67 ($J_{2b,2a}$ = 16.7) ($J_{2b,3}$ = 3.0)	1,3,8w,9,10,12,13,24
3	51.83	3-H	4.98 ($J_{3,2a}$ = 8.0) ($J_{3,2b}$ = 3.0)	1,2,5,8,24
4				
5	169.54			
6	100.35			
7		7-N		
8	156.07			1
9	100.38			
10	148.27			
11	135.15			
12	145.89			
13	137.26			
14	120.23	14-H	7.56 ($J_{14,16}$ = 1.0)	5,6,8w,9w,23

15	112.68			
16	131.85	16-H	8.22 ($J_{16,14}$ = 1.0)	15,18,23
17		17-NH		
18	137.20			
19	112.82	19-H	7.47 ($J_{19,20}$ = 7.9)	21,23
20	123.83	20-H	7.25 ($J_{20,19}$ = 7.9) ($J_{20,21}$ = 6.2)	18,22
21	122.04	21-H	7.23 ($J_{21,20}$ = 6.2) ($J_{21,22}$ = 7.8)	19,23
22	119.31	22-H	7.92 ($J_{22,21}$ = 7.8)	15,18,20
23	128.48			
24	171.23			
25		25-NH		
26	52.25	26-H	4.43 ($J_{26,27a}$ = 10.0) ($J_{26,27b}$ = 5.0)	24,27,28,31
27	31.34	27-H_a	1.73 ($J_{27a,27b}$ = 18.0) ($J_{27a,26}$ = 10.0)	26,28,29,30,31
		27-H_b	1.92 ($J_{27b,27a}$ = 18.0) ($J_{27b,26}$ = 5.0)	26,28,29,30,31
28	40.82	28-H	1.74	26,27,29,30
29	62.22	29-H_a	3.57	27,28,30
		29-H_b	3.61	27,28,30
30	63.68	30-H_a	3.57	27,28,29
		30-H_b	3.61	27,28,29
31	175.71			
32	60.84	32-H_3		11

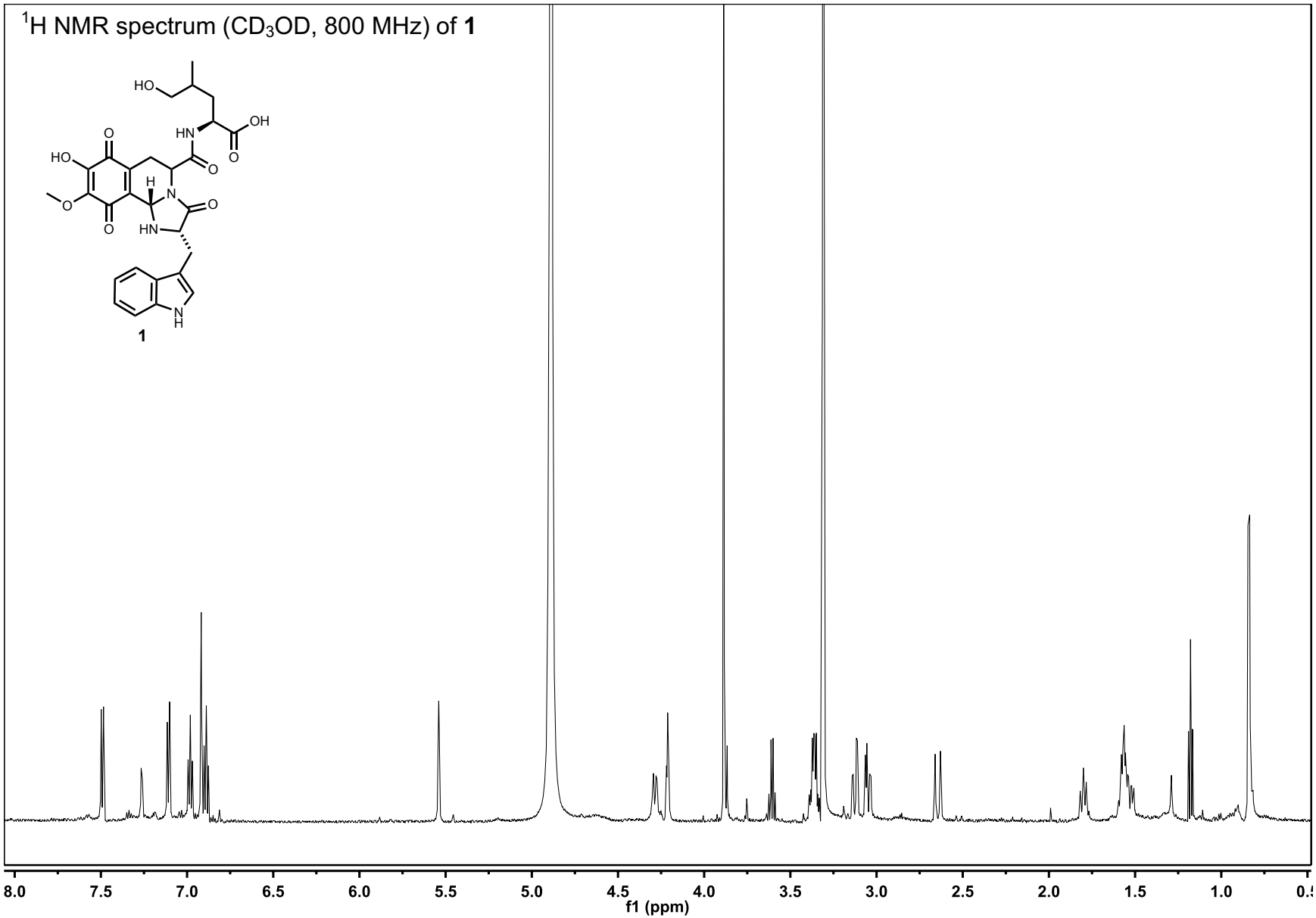
Table D8. ^{1}H (800 MHz) and ^{13}C (200 MHz) NMR spectroscopic data for compound **7** in methanol-d_4.

Chemical shifts were referenced to δ(C<u>H</u>D_2OD) = 3.31 ppm and δ(^{13}<u>C</u>H$D_2$$OD_3$) = 49.00.$^{13}C$ chemical shifts were determined via HMBC, HSQC and direct observation ^{13}C spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

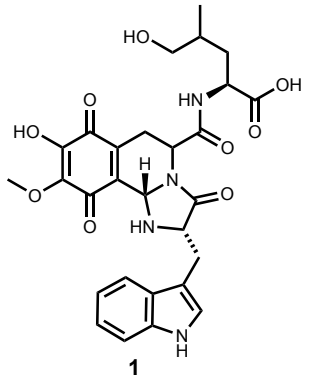


No.	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1	134.85			
2	26.31	2-H_a	0.05 ($J_{2a,2b}$ = 18.5) ($J_{2a,3}$ = 11.7)	1,3,4,10,11,14
		2-H_b	2.28 ($J_{2b,2a}$ = 18.5) ($J_{2b,3}$ = 5.0)	1,3,4,10,11,14
3	54.57	3-H	3.82 ($J_{3,2a}$ = 11.7) ($J_{3,2b}$ = 5.0)	1,2,4,7,9
4	168.39			
5		5-NH		
6	56.53	6-H	4.33 ($J_{6,16a}$ = 3.5) ($J_{6,16b}$ = 3.5) ($J_{6,3}$ = 1.0)	4,7,15
7	166.43			
8				
9	39.03	9-H_a	3.31 ($J_{9a,9b}$ = 18.2)	1,3,10,11,14
		9-H_b	5.03 ($J_{9b,9a}$ = 18.2)	1,3,10,11,14
10	135.77			
11	182.57			
12	139.18			
13	137.21			
14	183.45			
15	60.43	15-H_3	3.88	12
16	31.85	16-H_a	3.01 ($J_{16a,16b}$ = 14.7) ($J_{16a,6}$ = 3.5)	6,7,17,18,25
		16-H_b	3.57 ($J_{16b,16a}$ = 14.7) ($J_{16b,6}$ = 3.5)	6,7,17,18,25
17	108.74			
18	126.20	18-H	7.04	16,17,20,25
19				

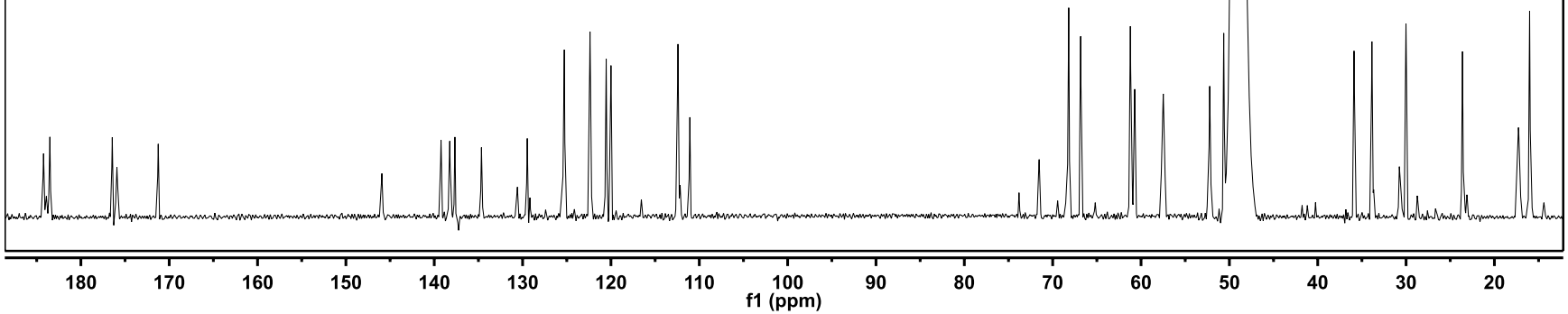
20	137.21			
21	112.62	21-H	7.04 ($J_{20,21}$ = 7.5)	23,25
22	122.59	22-H	6.88 ($J_{22,21}$ = 7.5) ($J_{22,23}$ = 7.5)	20,24
23	119.53	23-H	6.71 ($J_{23,22}$ = 7.5) ($J_{23,24}$ = 7.8)	21,25
24	119.85	24-H	7.39 ($J_{24,23}$ = 7.8)	17,20,22
25	128.76			

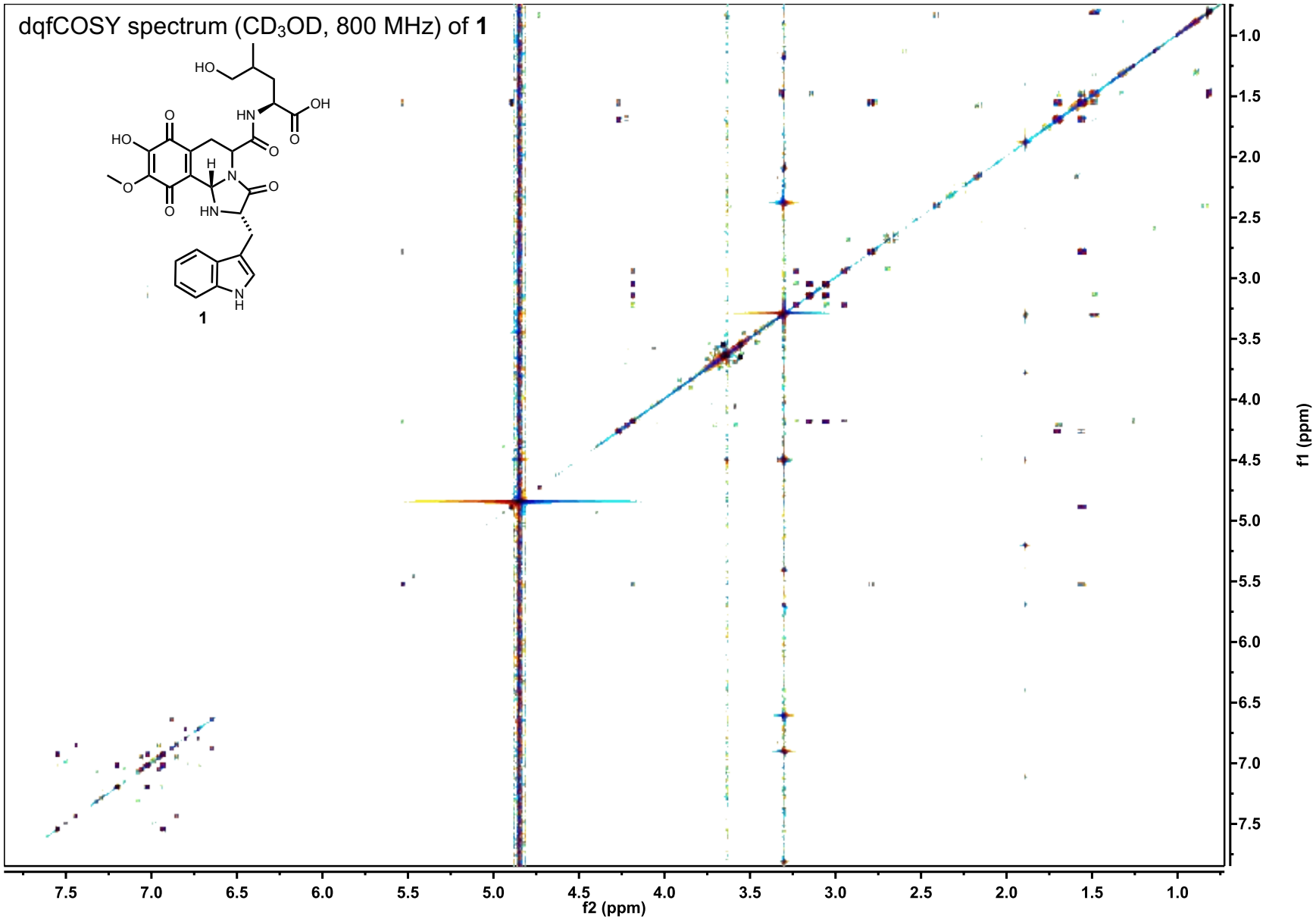

^{1}H NMR spectrum (CD_3OD, 800 MHz) of **1**
HO
HN
OH
O
O
HO
O
H
N
O
O
HN
O
N
H
1
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)

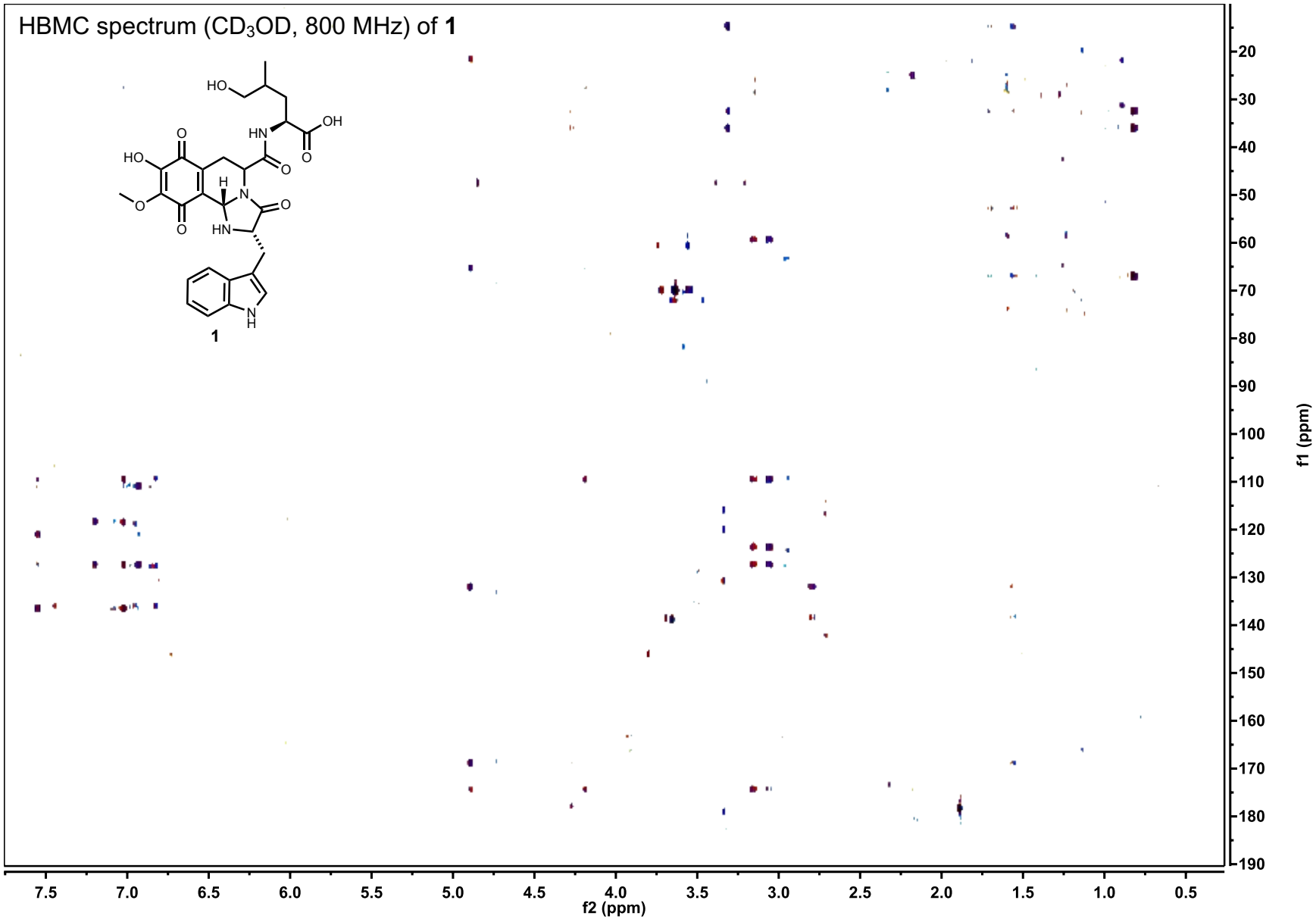
^{13}C NMR spectrum (CD_3OD, 800 MHz) of **1**

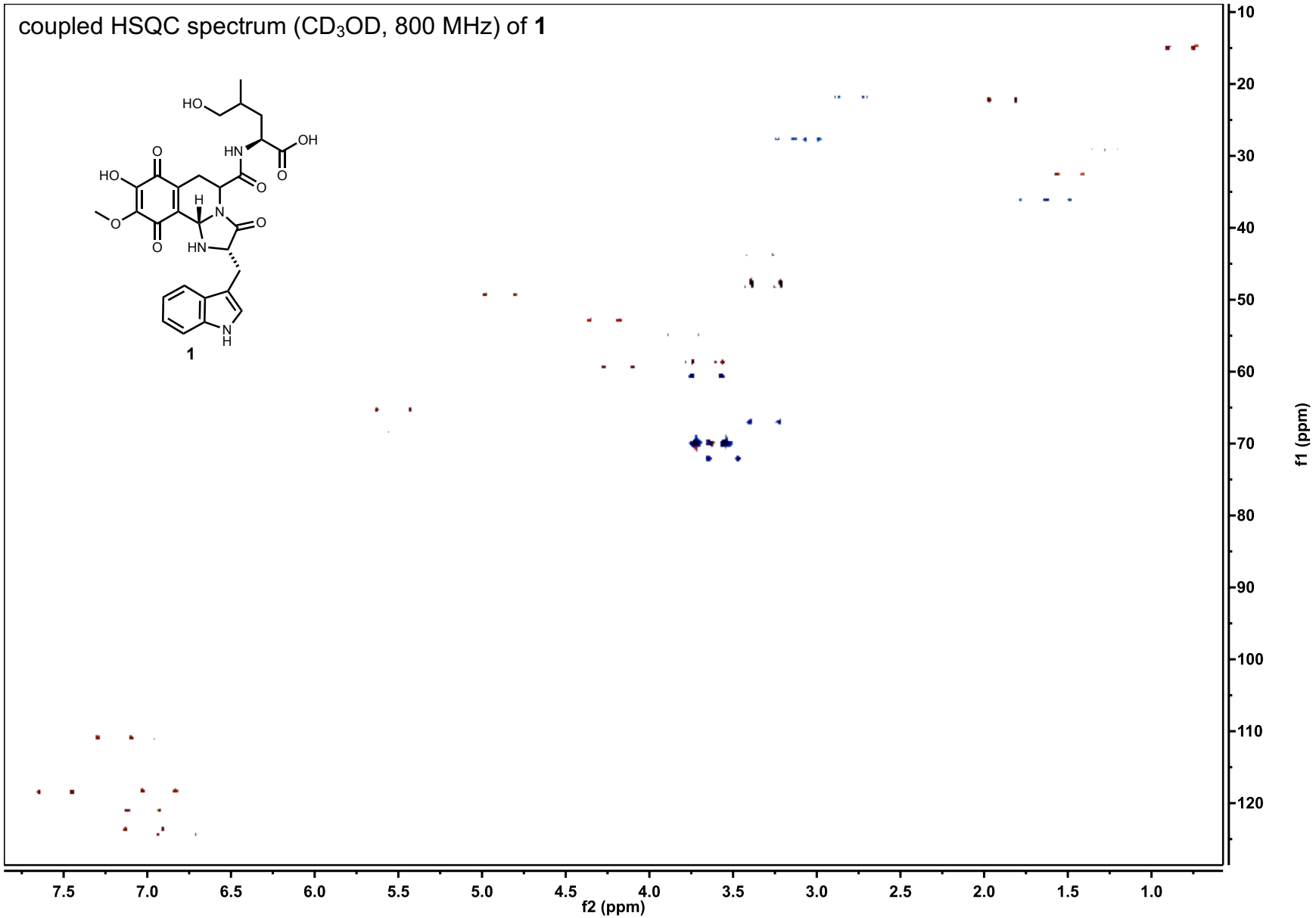


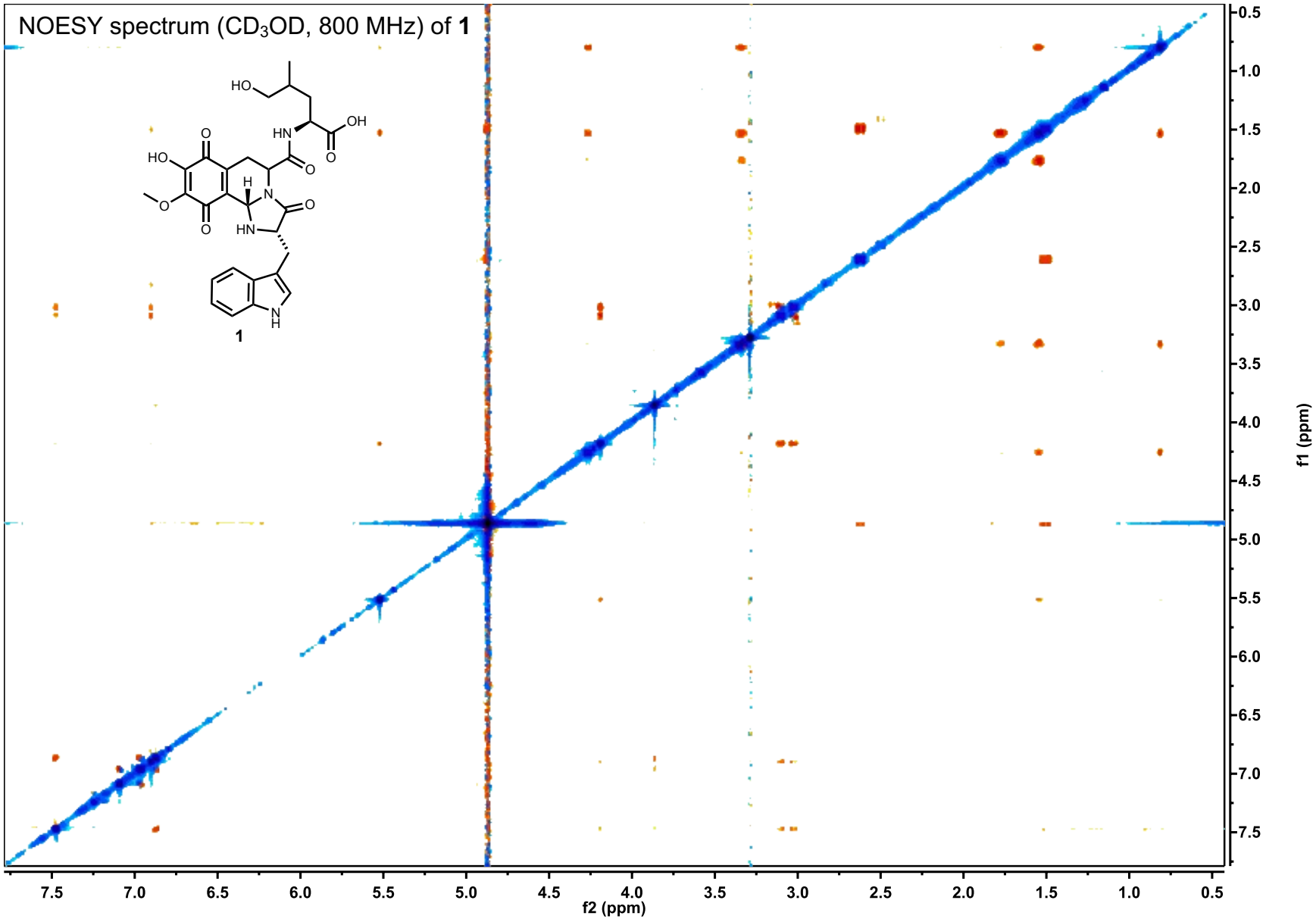
1



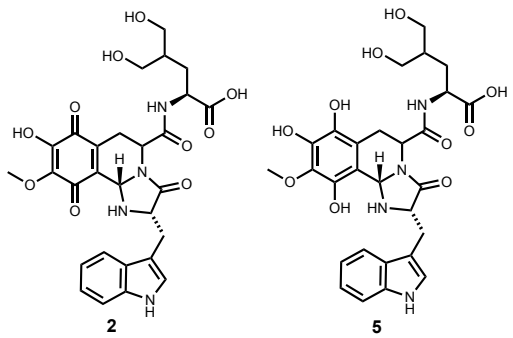

dqfCOSY spectrum (CD_3OD, 800 MHz) of **1**
HO
HN
OH
O
HO
H
N
O
O
HN
O
O
N
H
1
f2 (ppm)
f1 (ppm)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0


HBMC spectrum (CD3OD, 800 MHz) of 1
HO
HN
OH
O
O
HO
O
O
H
N
O
O
HN
N
H
1
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
f1 (ppm)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f2 (ppm)


coupled HSQC spectrum (CD_3OD, 800 MHz) of **1**
HO
HN
OH
O
HO
H
N
O
O
HN
O
O
N
H
1
10
20
30
40
50
60
70
80
90
100
110
120
f1 (ppm)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f2 (ppm)


NOESY spectrum (CD3OD, 800 MHz) of 1
HO
OH
HN
O
HO
O
H
N
O
O
O
HN
N
H
1
f2 (ppm)
f1 (ppm)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

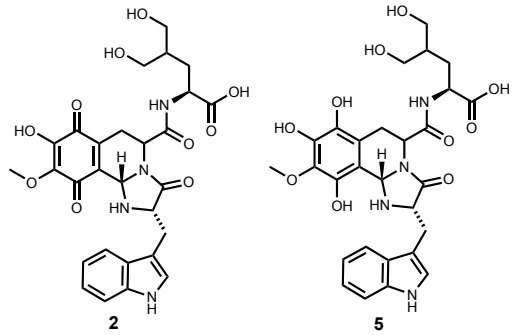
^{1}H NMR spectrum (CD_3OD, 800 MHz) of a mixture of **2** and **5**.



2

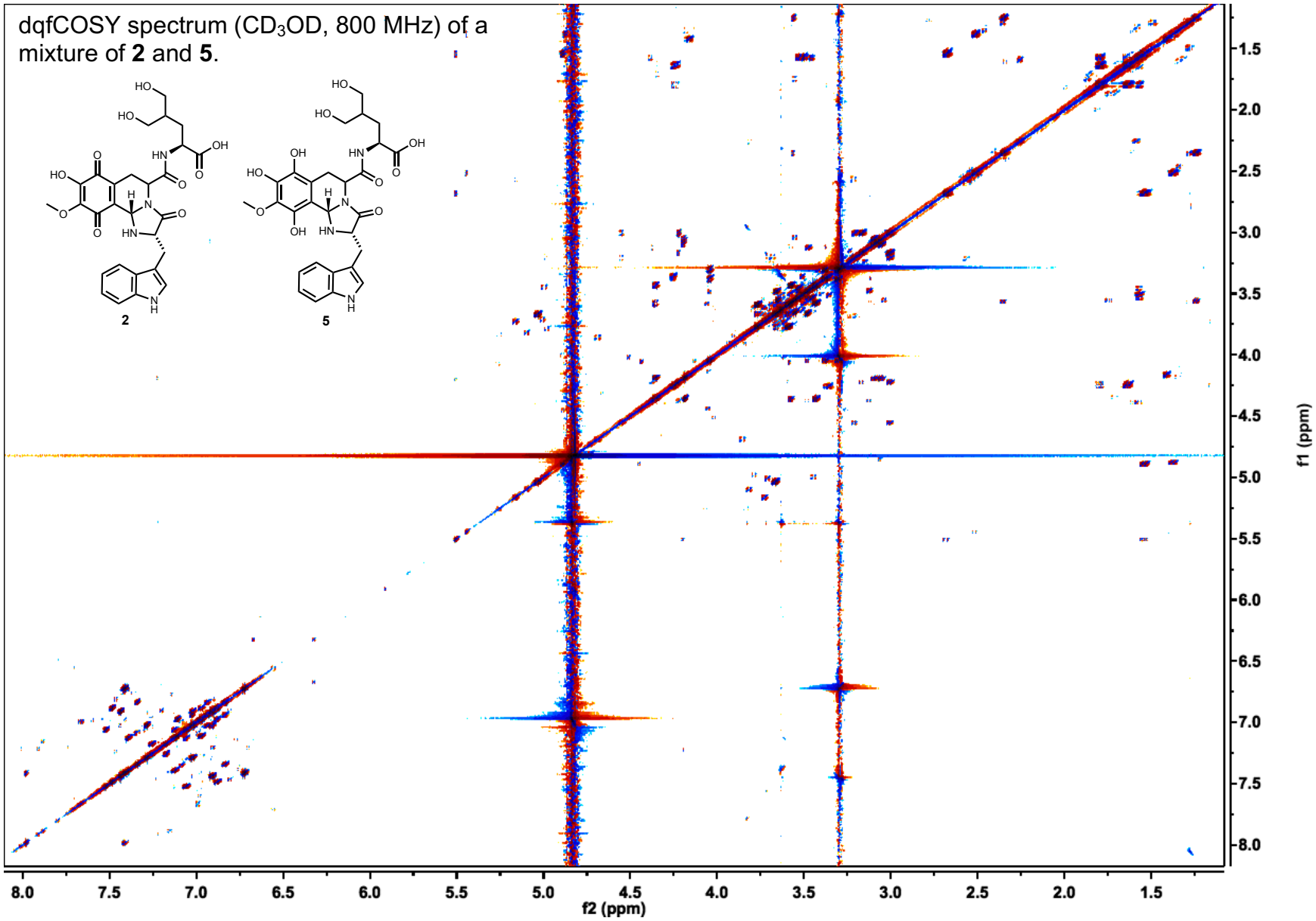
5

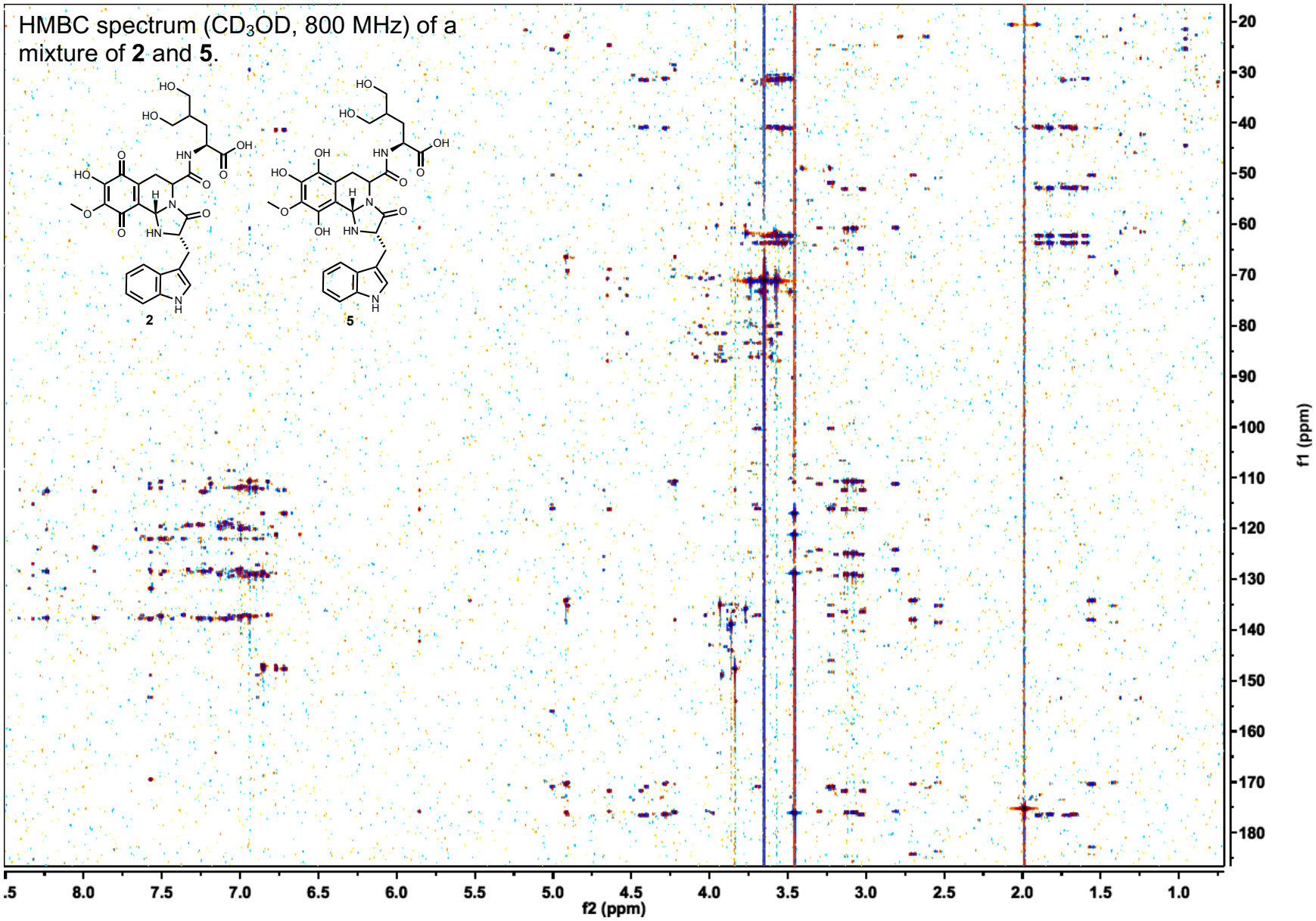
^{13}C NMR spectrum (CD_3OD, 800 MHz) of a mixture of **2** and **5**.

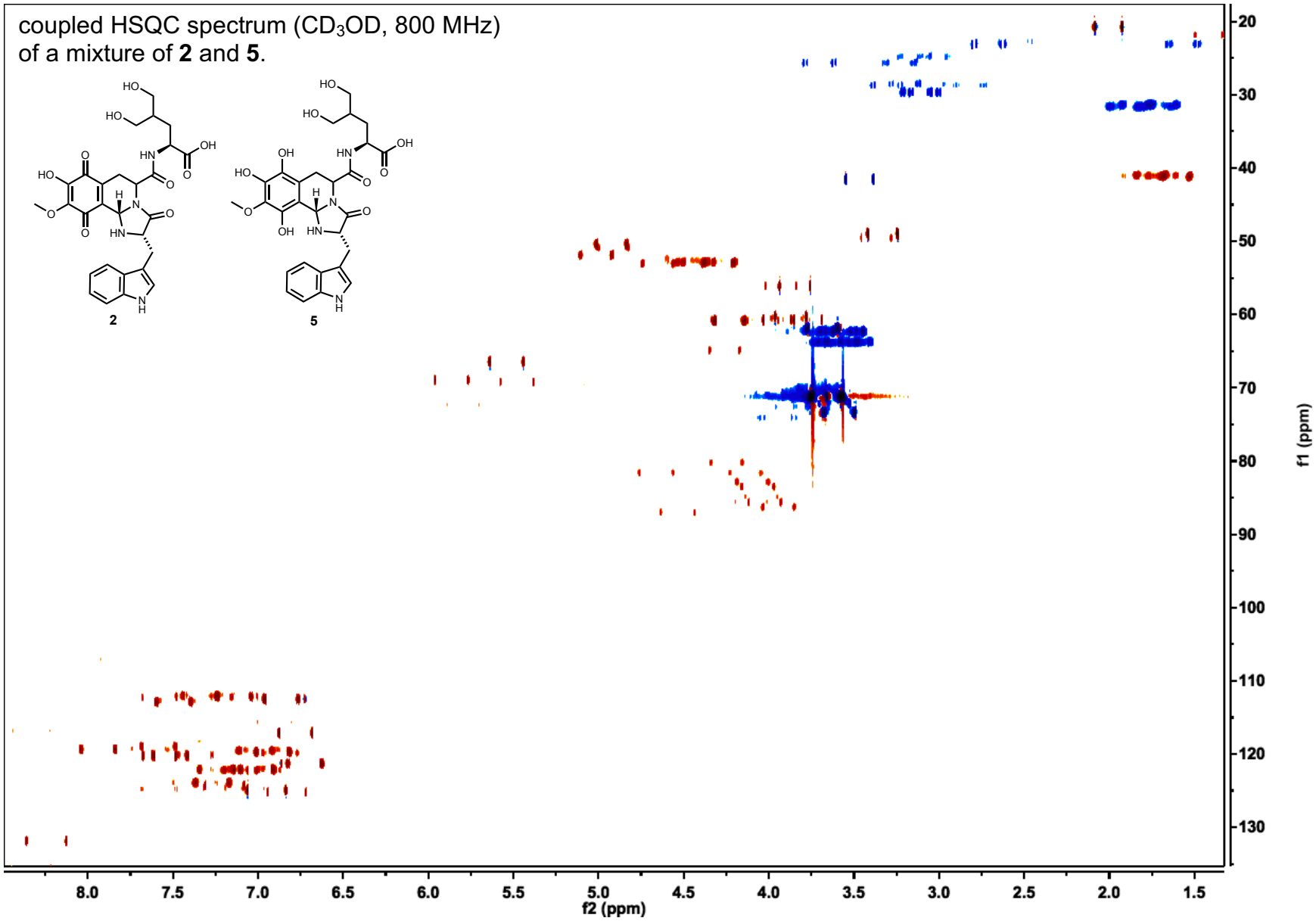


2

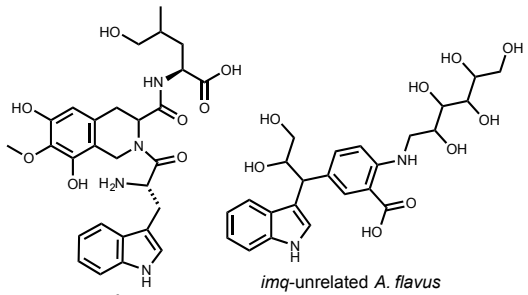
5


dqfCOSY spectrum (CD_3OD, 800 MHz) of a mixture of **2** and **5**.
2
5
f2 (ppm)
f1 (ppm)


HMBC spectrum (CD_3OD, 800 MHz) of a mixture of **2** and **5**.
2
5
f2 (ppm)
f1 (ppm)

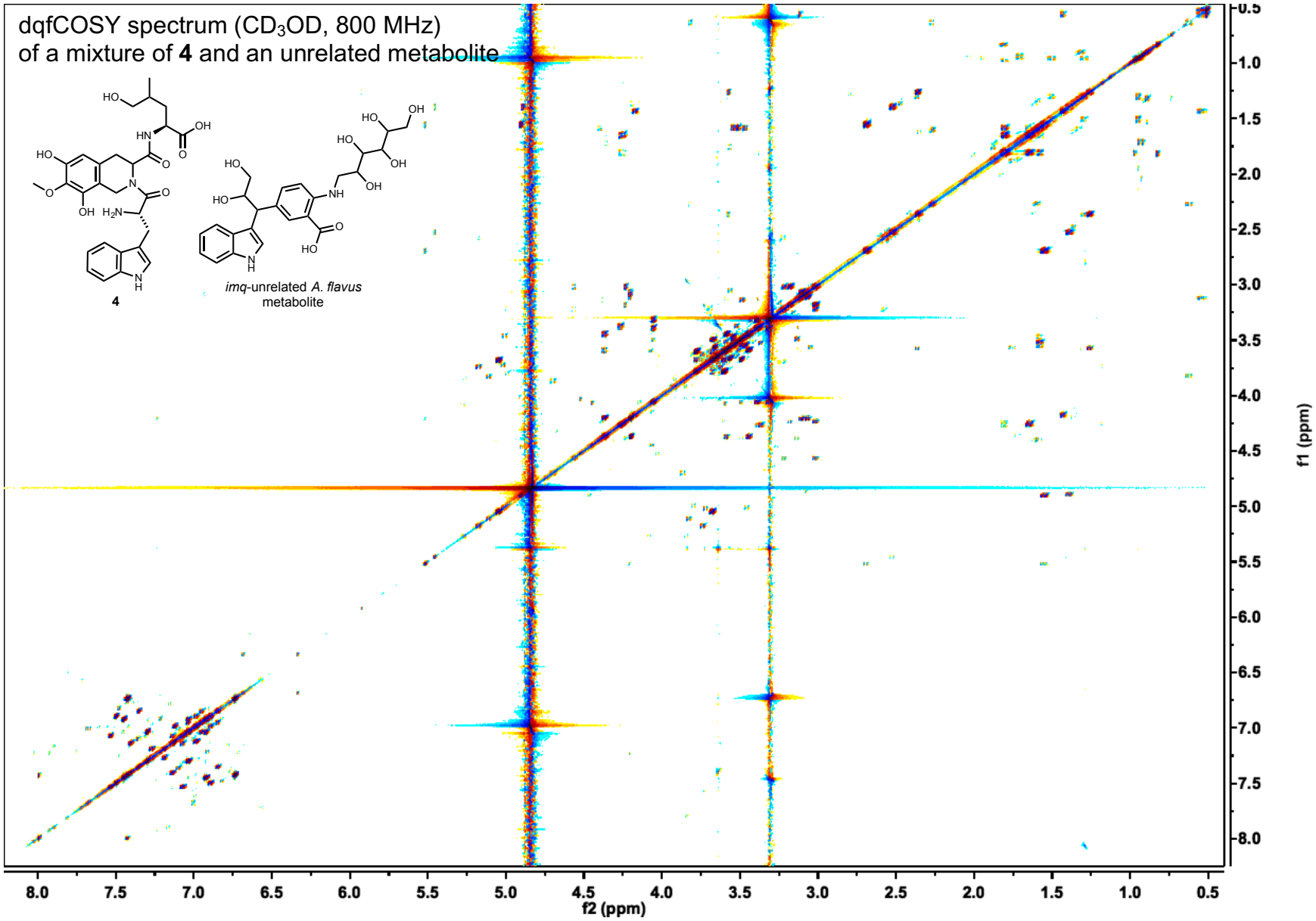

coupled HSQC spectrum (CD_3OD, 800 MHz)
of a mixture of **2** and **5**.
2
5
f2 (ppm)
f1 (ppm)

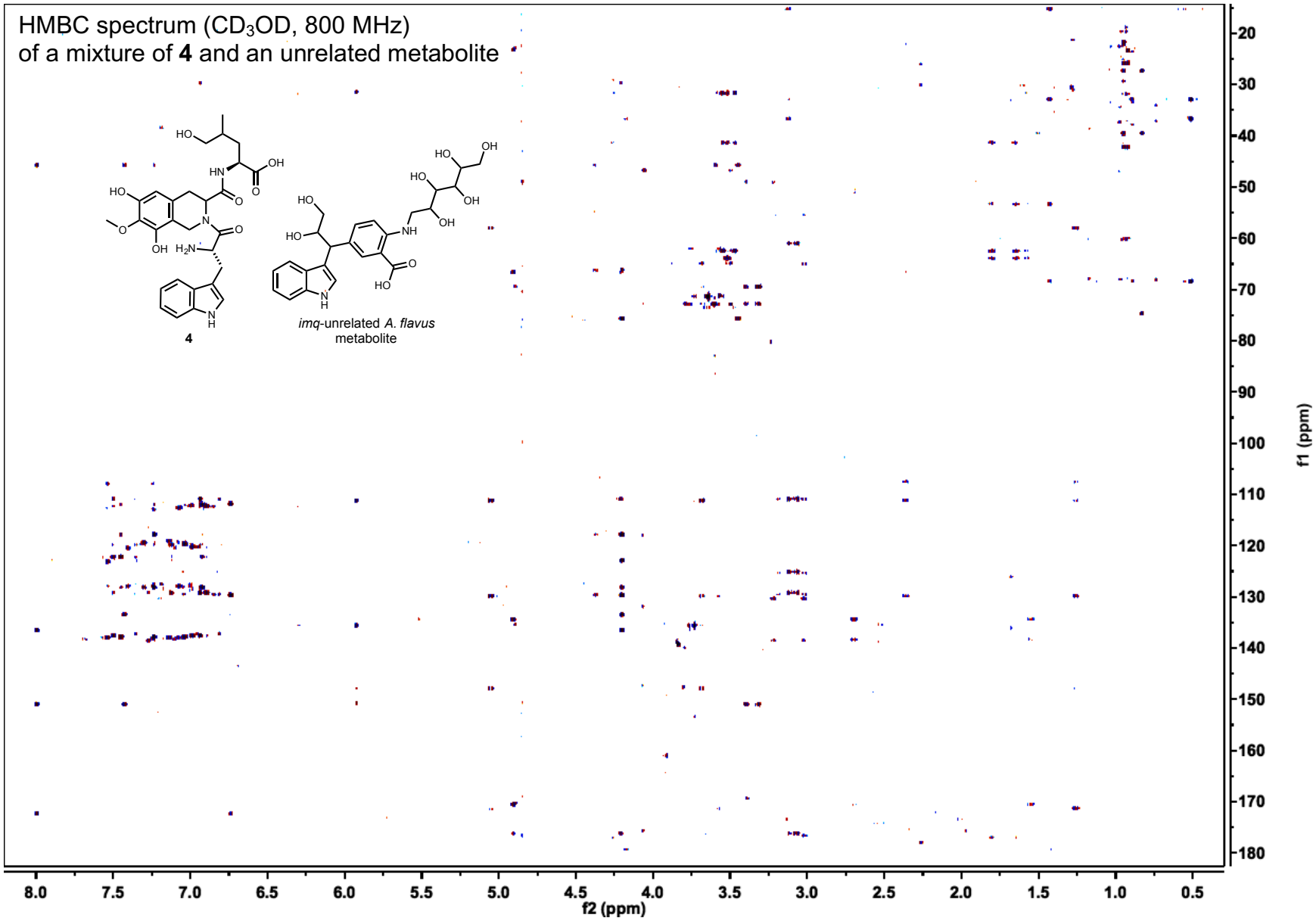
^{1}H NMR spectrum (CD_3OD, 800 MHz)
of a mixture of **4** and an unrelated metabolite

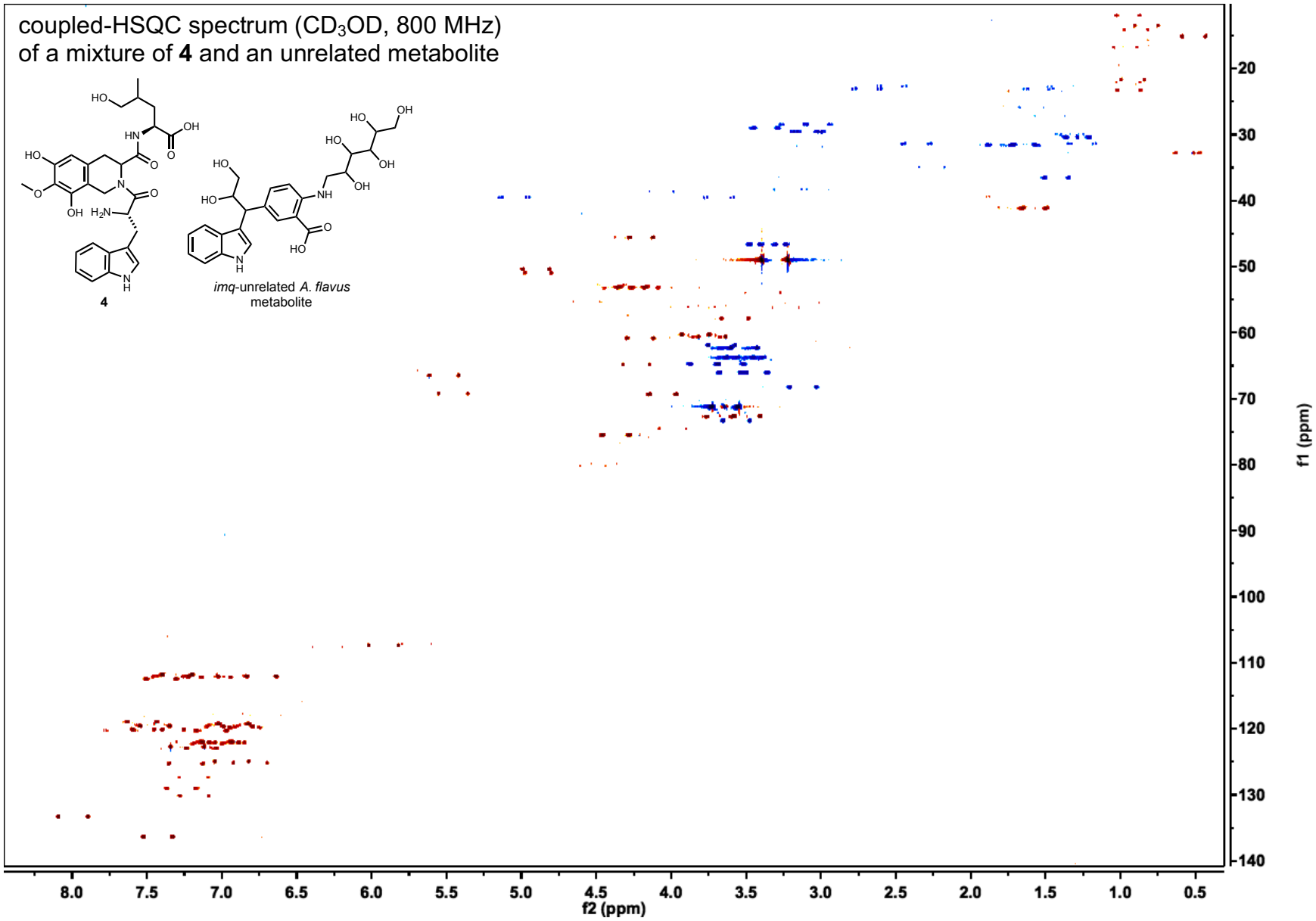


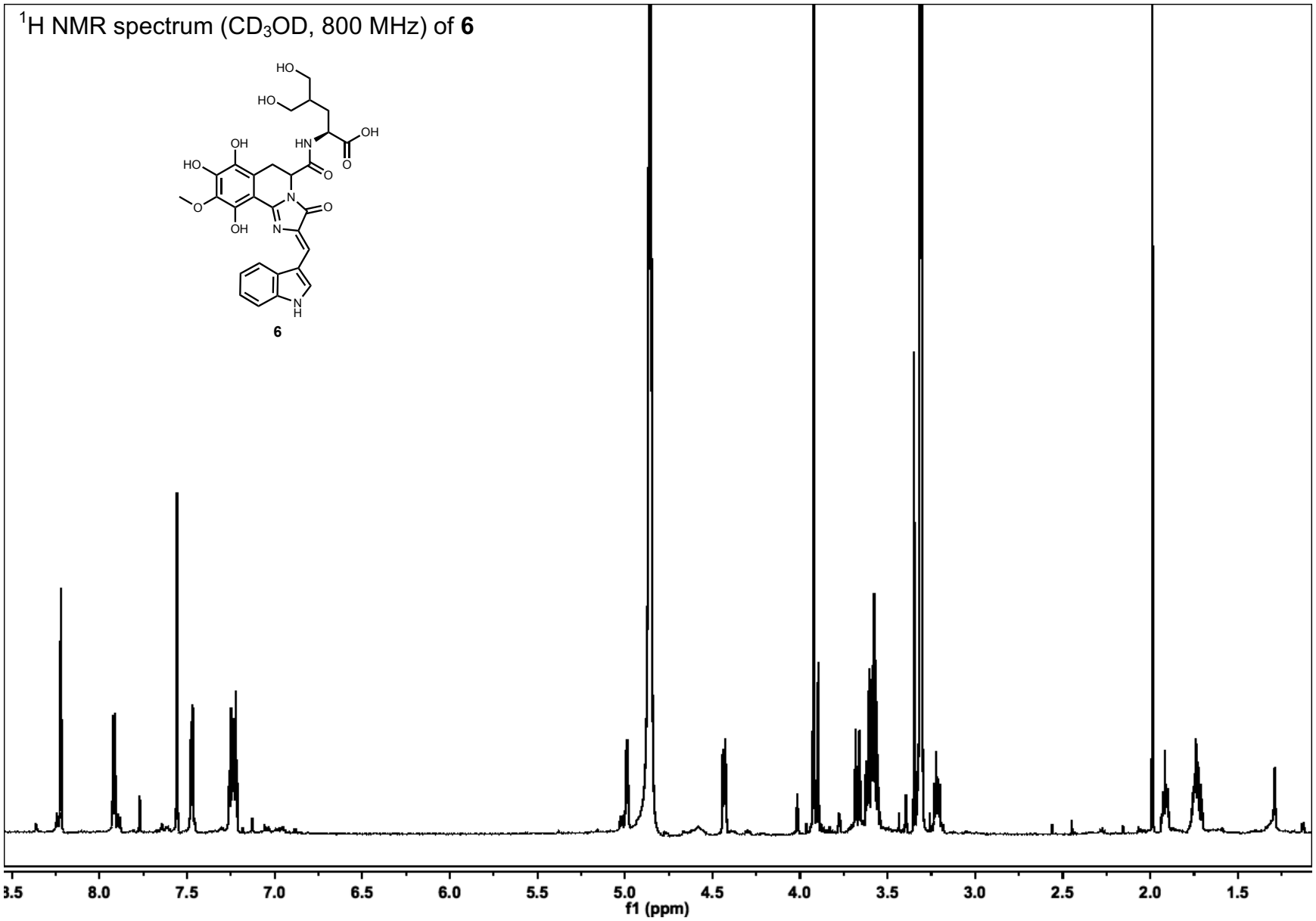
4

imq-unrelated *A. flavus* metabolite

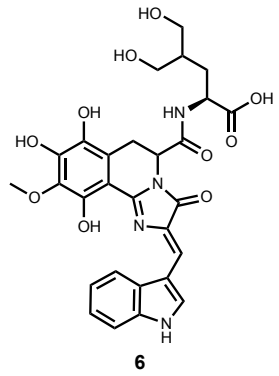

dqfCOSY spectrum (CD_3OD, 800 MHz)
of a mixture of **4** and an unrelated metabolite
4
imq-unrelated *A. flavus*
metabolite
f2 (ppm)
f1 (ppm)


HMBC spectrum (CD_3OD, 800 MHz)
of a mixture of **4** and an unrelated metabolite
4
imq-unrelated *A. flavus* metabolite
f2 (ppm)
f1 (ppm)


coupled-HSQC spectrum (CD_3OD, 800 MHz)
of a mixture of **4** and an unrelated metabolite
4
imq-unrelated *A. flavus*
metabolite
f2 (ppm)
f1 (ppm)


^{1}H NMR spectrum (CD_3OD, 800 MHz) of **6**
HO
HO
OH
HN
OH
O
HO
O
N
O
O
OH
N
N
H
6
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
f1 (ppm)

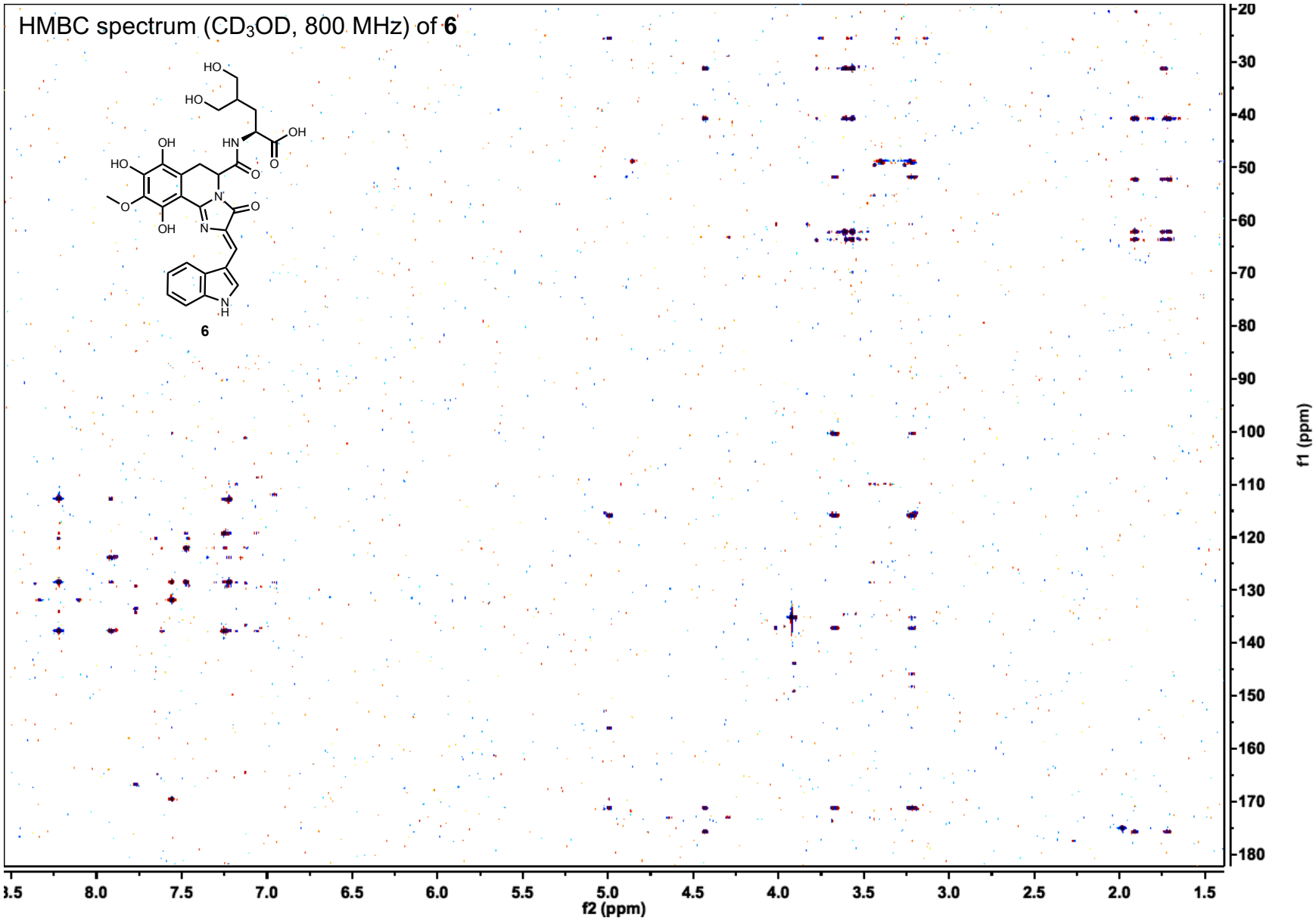
dqfCOSY spectrum (CD_3OD, 800 MHz) of **6**



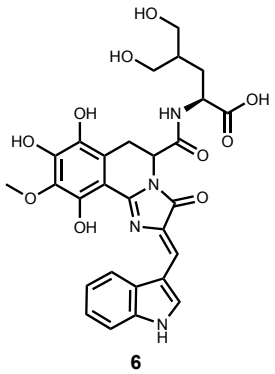
6

f2 (ppm)

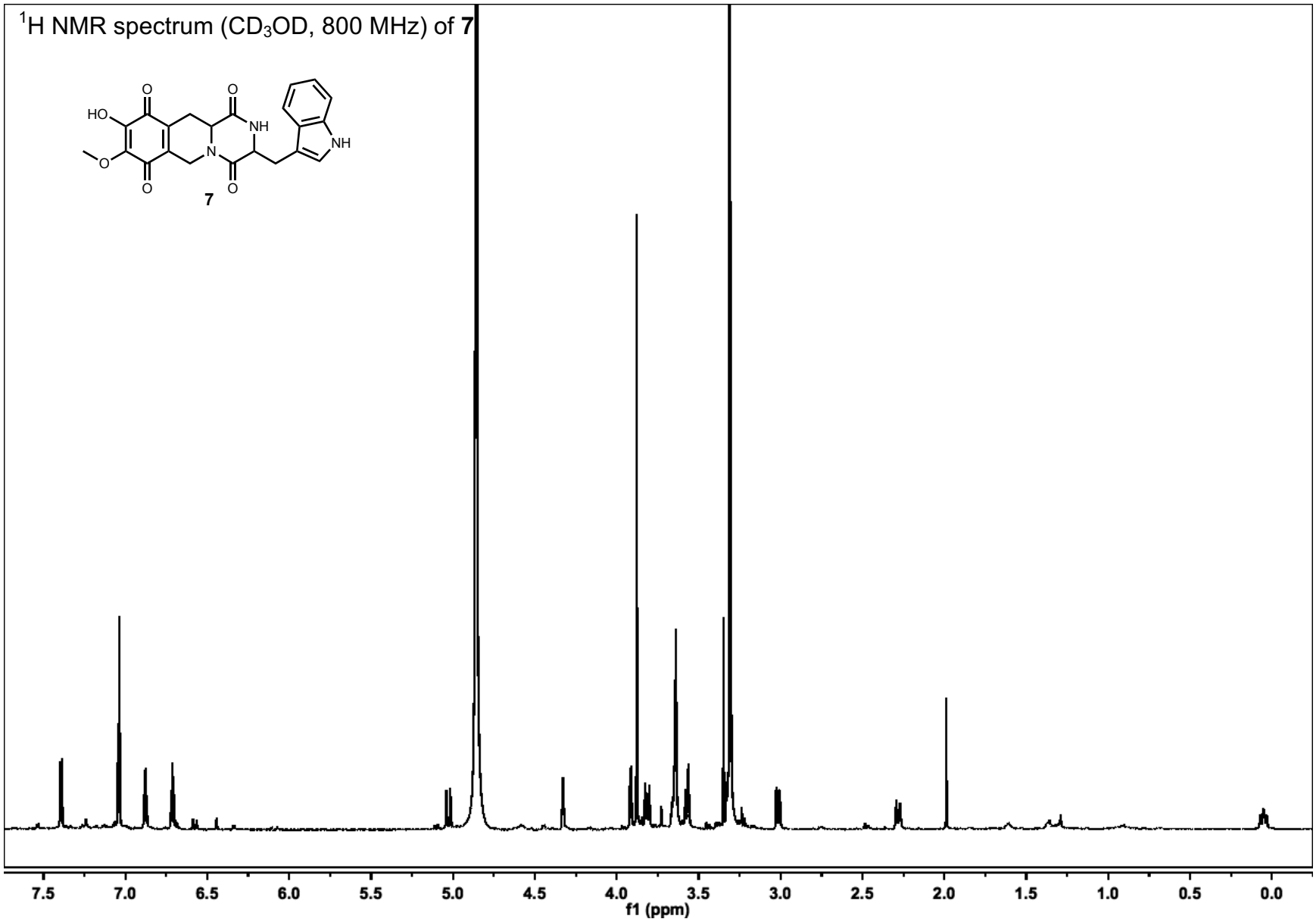
f1 (ppm)

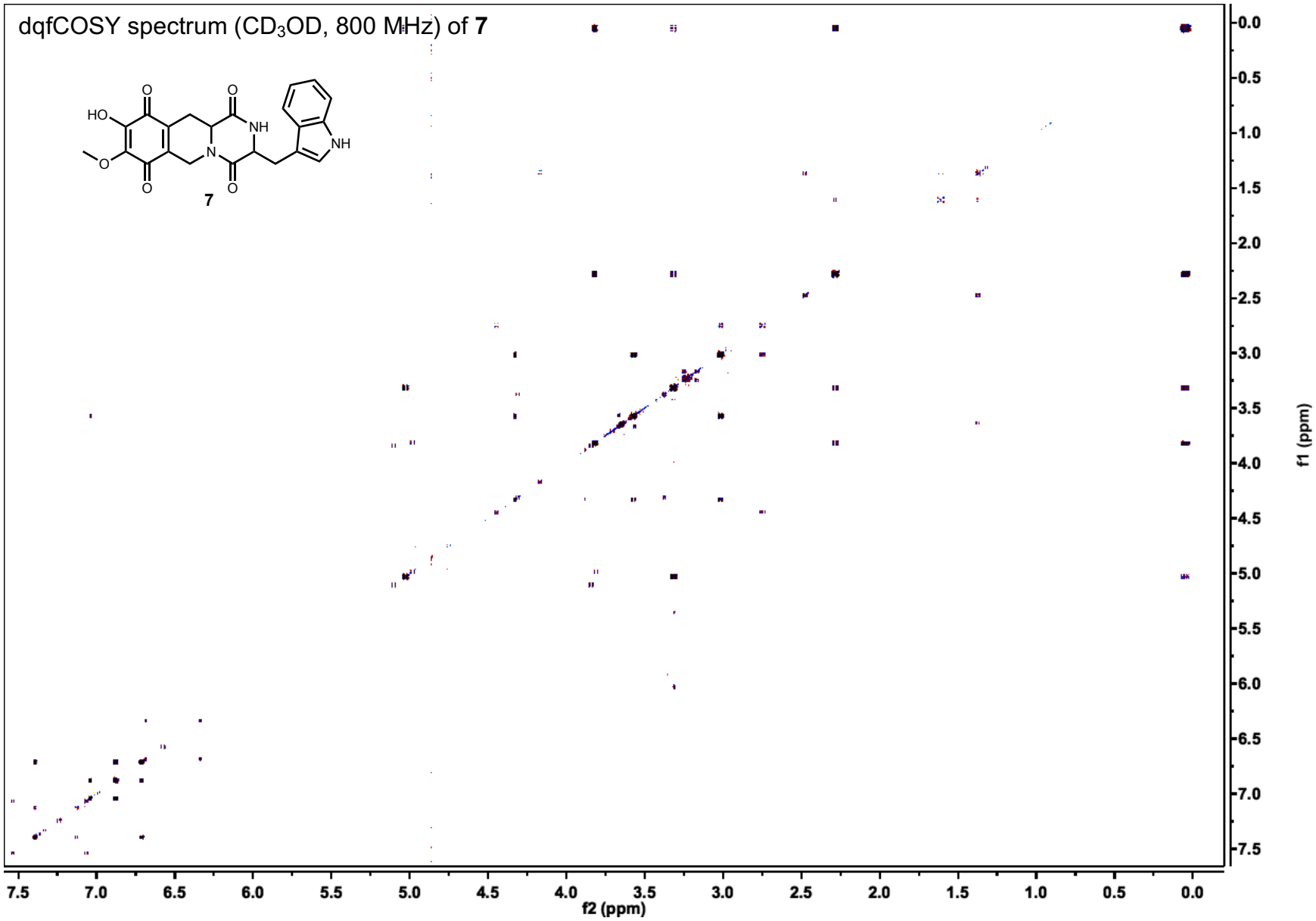

HMBC spectrum (CD3OD, 800 MHz) of 6
HO
HO
OH
HN
OH
HO
O
O
N
O
OH
N
N
H
6
f2 (ppm)
f1 (ppm)

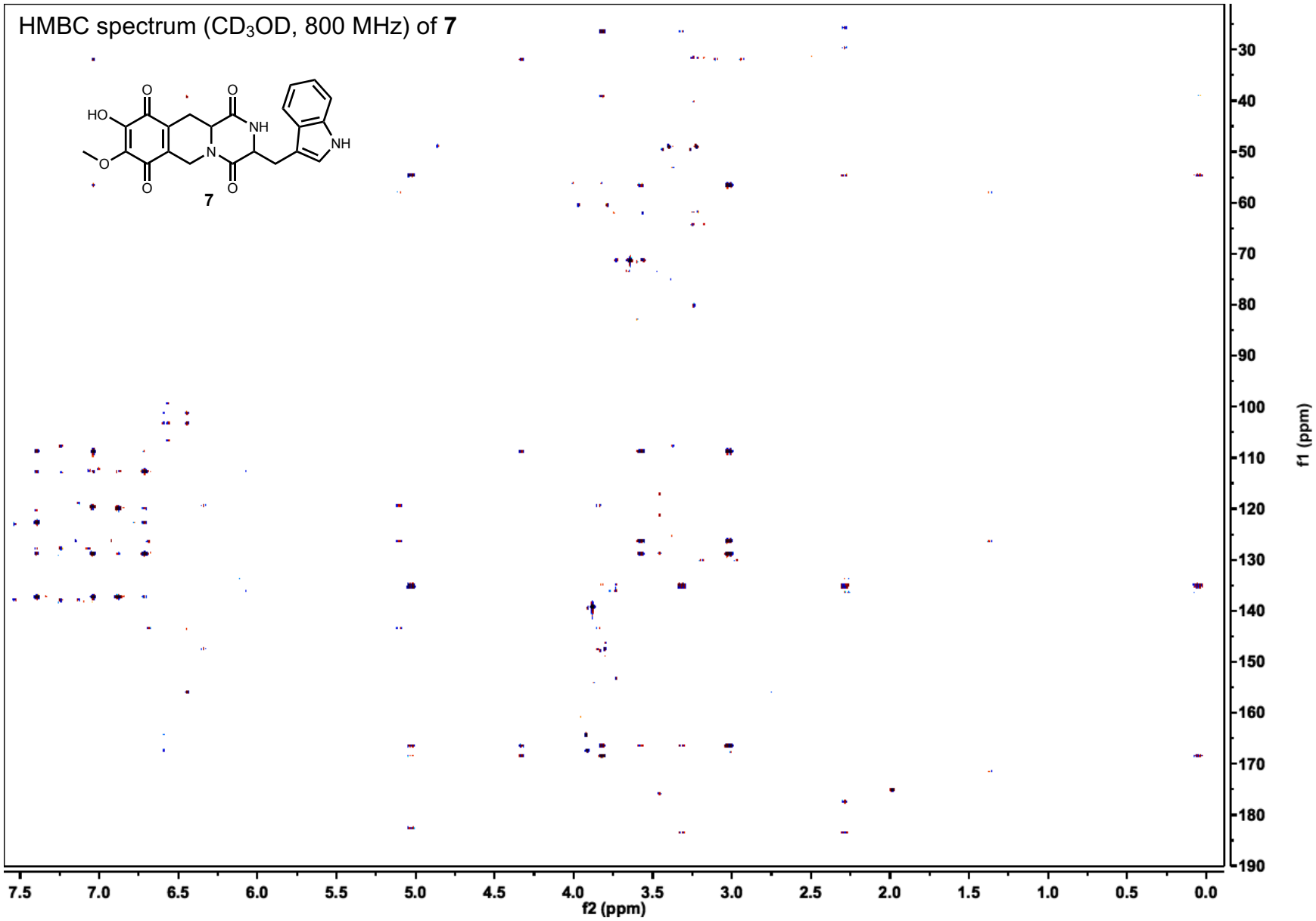
coupled HSQC spectrum (CD_3OD, 800 MHz) of **6**

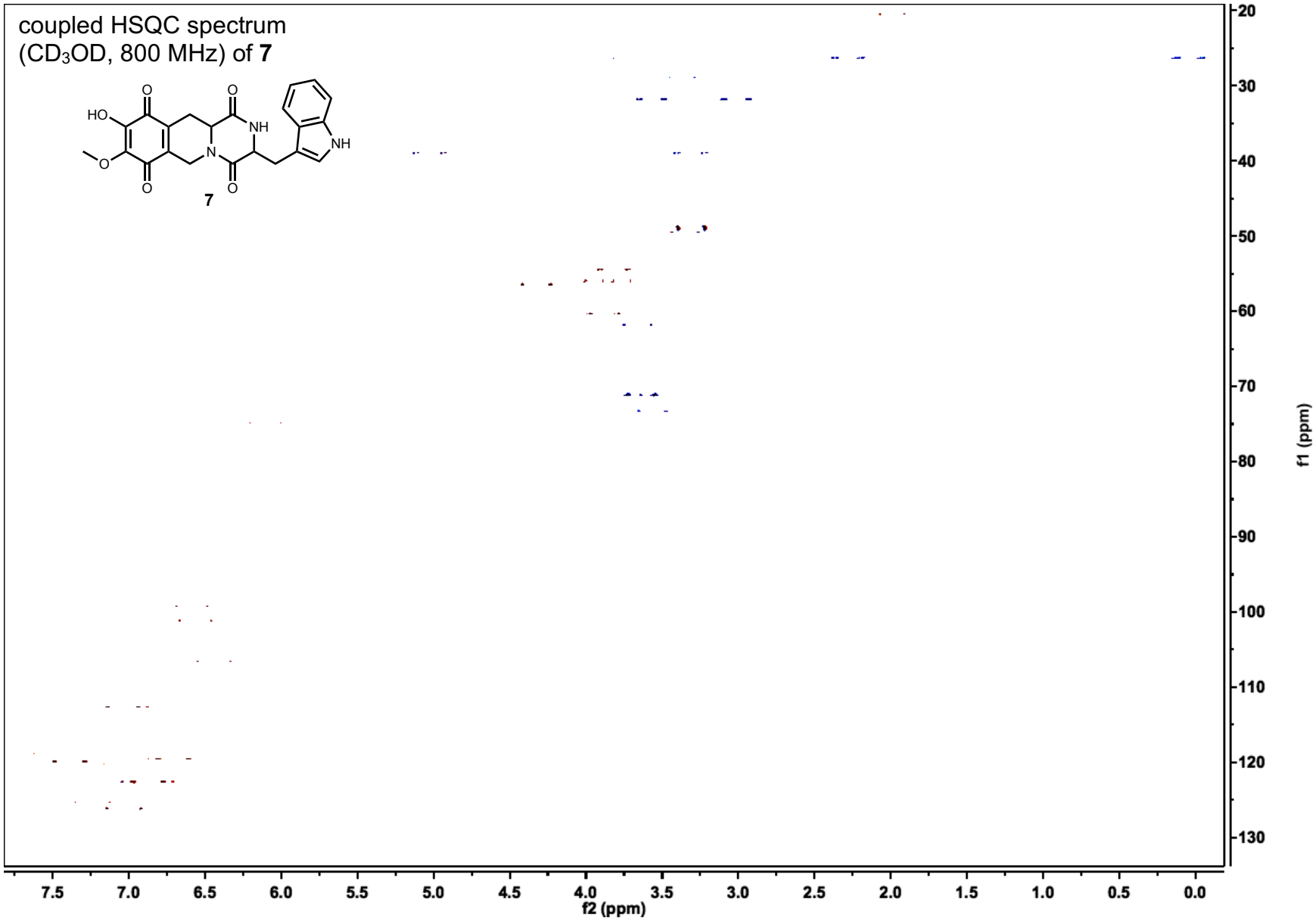


6


^{1}H NMR spectrum (CD_3OD, 800 MHz) of **7**
HO
O
O
NH
N
NH
O
O
O
O
7
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


dqfCOSY spectrum (CD3OD, 800 MHz) of 7
HO
O
O
NH
N
NH
O
O
O
7
f2 (ppm)
f1 (ppm)


HMBC spectrum (CD3OD, 800 MHz) of 7
HO
O
O
NH
N
NH
O
O
O
7
f2 (ppm)
f1 (ppm)


coupled HSQC spectrum
(CD3OD, 800 MHz) of 7
HO
O
O
NH
N
NH
O
O
O
7
f2 (ppm)
f1 (ppm)

REFERENCES

1. Shimizu, K. & Keller, N. P. Genetic involvement of a cAMP-dependent protein kinase in a G protein signaling pathway regulating morphological and chemical transitions in Aspergillus nidulans. *Genetics* **157,** 591–600, (2001).

2. He, Z.-M., Price, M. S., OBrian, G. R., Georgianna, D. R. & Payne, G. A. Improved protocols for functional analysis in the pathogenic fungus Aspergillus flavus. *BMC Microbiol* **7,** 104, (2007).

APPENDIX E

BIOSYNTHESIS OF THE FUMIZINONES: SECONDARY PRODUCTS OF THE FSQ CLUSTER IN ASPERGILLUS FUMIGATUS

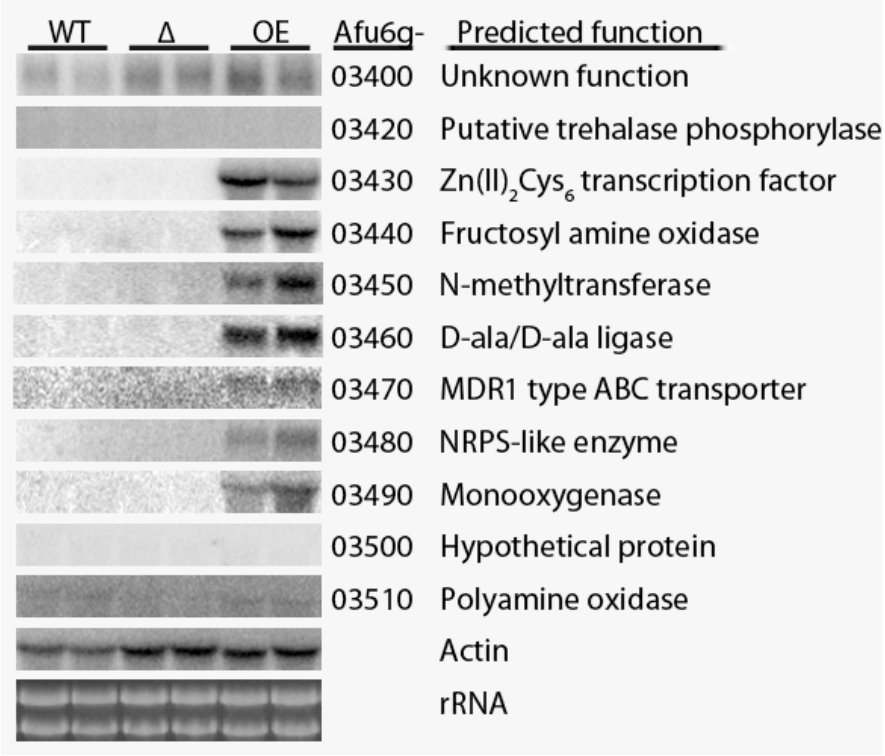


Figure E1. Northern blot analysis of WT (Af293) left, Δ*fsqA* middle, and OE::*fsqA,* right. Overexpression of *fsqA* causes specific up-regulation of AFUA6g_03430 – AFUA6g_03490, defining the boundaries of the *fsq* cluster.

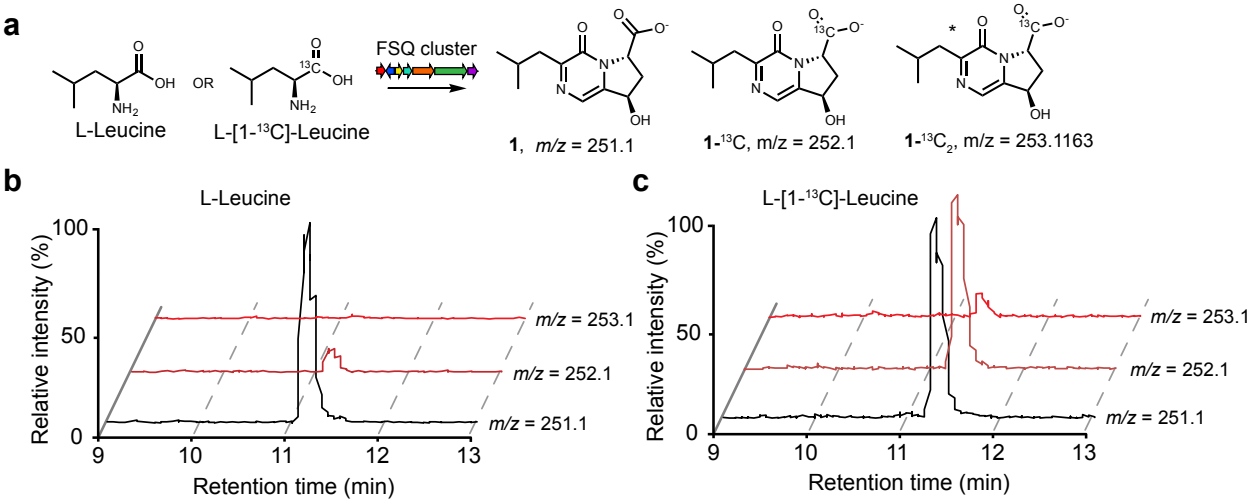


Figure E2. L-Leucine is incorporated into the fumizinones. (**a**) Feeding experiments with L-Leucine or L-[1-^{13}C]-Leucine result in a distribution of isotopes for **1**. (**b**) LC-MS extracted-ion-chromatograms from whole metabolome analysis of OE::*fsqA* fed with unlabeled control, L-Leucine. (**c**) LC-MS extracted-ion-chromatograms from whole metabolome analysis of OE::*fsqA* fed with unlabeled control, L-Leucine. * normally isotopic abundance of 13C results in higher M+1 for L-Leucine control and higher M+2 for L-[1-^{13}C]-Leucine.

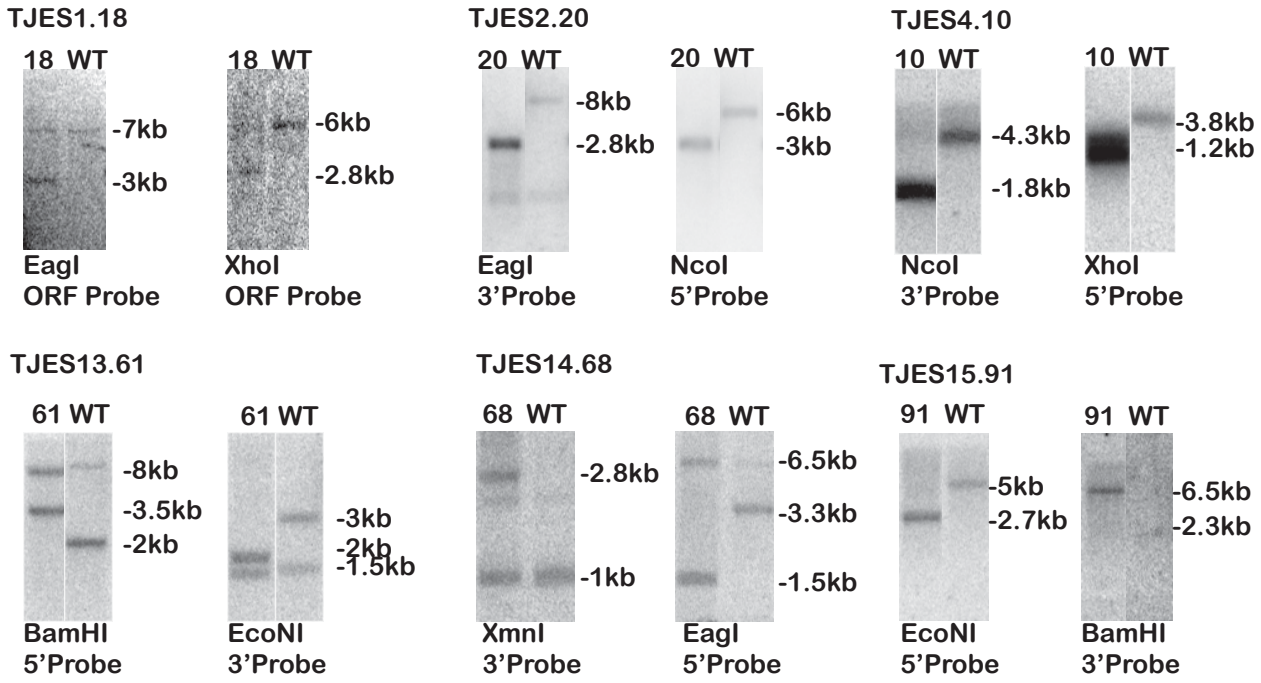


Figure E3. Southern analysis confirmation of all mutants used in this study. Mutant (number) lanes are on the left of each image and the parental strain ("WT") is shown on the right. Expected band sizes correlating with those seen in the images are marked accordingly. Under each image is the enzymes used in the restriction digest as well as the nucleic acid probe used.

Strains, media, and growth conditions. The fungal strains used in this study are listed in **Table E2**. All strains were grown at 37 °C on glucose minimal medium (GMM)[1] and, when appropriate, were supplemented with 0.56 g/L uracil, 1.26 g/L uridine, 1.0 g/L arginine and maintained as glycerol stocks at -80 ºC. *Escherichia coli* strain DH5α was propagated in LB medium with appropriate antibiotics for plasmid DNA.

Gene cloning, plasmid construction, and genetic manipulations. The plasmids utilized in this study are listed in **Table E2**. The oligonucleotide sequences for PCR primers are given in **Table E3**. PCR amplification was carried out on a C1000^{TM} Thermal Cycler (Bio-Rad). For creation of the *fsqA* overexpression (OE) strain (TJES1.18) the OE cassette was constructed by amplifying the *fsqA* open reading frame (ORF) from *Af293* genomic DNA using primers Afu6g03430_NcoI_For and Afu6g03430_NotI_Rev, which introduced a 5' NcoI restriction site and a 3' NotI restriction site. This PCR product was purified with a QIAquick gel extraction kit (Qiagen), quantified, and digested with the appropriate restriction enzymes before being cloned into pJMP8.12[2], resulting in plasmid pJES 1.2. Subsequently, *A. fumigatus argB* was transferred from pJMP4.1[3] by digesting with EcoRI and introduced into pJES1.2, resulting in plasmid pJES2.7 (*A. fumigatus argB::gpdA(p)::fsqA*). All PCR steps were carried out with Pfu Ultra II DNA Polymerases (Agilent) and all digestion reactions were carried out with NEB enzymes (New England BioLabs). Correctness of the inserted DNA was then confirmed by sequencing.

Construction of the *fsq* gene knock-out cassettes were generated using standard double-joint PCR procedures[4]. Briefly, *A. parasiticus pyrG* (*A.ppyrG)* was amplified from pJMP9.1. Then, an approximately 1 kb fragment upstream and downstream of each *fsq*

gene was amplified from genomic DNA of *A. fumigatus* Af293 using designated primers, respectively. These three amplified PCR products were purified with a QIAquick gel extraction kit, quantified, and fused using standard double-joint PCR procedures. The final PCR product was amplified with the bottom primer pairs, *gene*_5'F_flank and _3'R_flank, confirmed with endonuclease digestion, and purified for fungal transformation. The first two rounds of PCR were done with Pfu Ultra II DNA Polymerases (Agilent), and the final PCR step used Expand long template PCR system (Roche) according to the manufacturer's instructions. Af293.6 (double auxotroph *A. fumigatus, pyrG*$^{-}$*, argB*$^{-}$) was used to make the *OE*::*fsqA, pyrG*$^{-}$ auxotroph, TJES1.18. This strain was used as the recipient host strain for subsequent deletions of *fsq* genes as well as the ectopic complementation of *pyrG1* with plasmid pJMP9.1 (containing *A.ppyrG*) resulting in prototrophic strain TJES3.1 (OE::*fsqA, A.ppyrG*). Similarly Δ*fsqA* auxotrophic strains (TJES2.20) were made by transforming Af293.6 with deletion cassettes and complemented with ectopically integrated pJMP4 to make Δ*fsqA* prototroph TJES8.2.

For the creation of deletion mutants in OE*::fsqA* TJES1.18 background, a deletion cassette of each cluster gene (*fsqB* to *fsqG*) was constructed by using double-joint PCR with *A. parasiticus pyrG* (*A.ppyrG*) for replacement of target genes. The *A.ppyrG* gene was amplified from the plasmid pJMP9.1 as template. Five µg of the double-joint PCR cassette were used to delete each *fsq* cluster gene by using TJES1.18 (*A.fargB::gpdA(p):: fsqA, pyrG1*) as the recipient host. After transformation, transformants were grown on minimal media plates without supplements for screening. All strains were verified by PCR and Southern blot analysis (**Figure E3**). Multiple

confirmed strains for each mutant were stored at -80 °C with 33 % glycerol for future use.

Nucleic acid analysis. Preparation of plasmids, restriction enzyme digestions, gel electrophoresis, blotting, probe preparation and hybridization were carried out by standard protocols. *Aspergillus* DNA was extracted using a previously described method[7]. Sequence data were analyzed using the LASERGENE software package from DNASTAR.

Northern analysis. 50 mL of liquid GMM[1] were inoculated with 1.0×10^6 spores (asexual) mL^{-1} of all appropriate strains in this study and incubated with shaking at 250 rpm at 25 °C. After 48 h, the mycelium was collected and total RNA was extracted by using Isol-RNA Lysis Reagent according to the manufacturer's instructions (5 Prime).

In order to determine the boundaries of the *fsq* gene cluster, gene fragments of potential cluster genes used as probes were amplified individually from Af293 genomic DNA with appropriate primers. WT, OE::*fsqA* (TJES3.1), and Δ*fsqA* (TJES8.2) were used for this experiment. About 40 µg of total RNA were used for RNA blot analysis. RNA blots were hybridized with designated DNA fragments. All experiments were performed in duplicate. Detection of signals was carried out with a Storm 860 phosphoimager (Molecular Dynamics).

Fermentation and metabolome extraction for comparative metabolomics by LC-MS. Preparation for LC-MS analysis: *A*. *fumigatus* strains were inoculated (1.0×10^6 spores mL^{-1}) into 50 mL GMM[1] in a 125 mL Erlenmeyer flask at 37 °C with shaking at 220 rpm. After 4 days, liquid fungal cultures including fungal tissue and media were

frozen using a dry ice acetone bath, and lyophilized. The lyophilized residues were extracted with 30 mL of MeOH for 1.5 h with vigorous stirring. Extracts were filtered over cotton, evaporated to dryness, and stored in 4 mL vials. Crude extracts were suspended in 0.5 mL of MeOH and centrifuged to remove insoluble materials, and the supernatant was subjected to HPLC-MS analysis.

Fermentation and metabolome extraction for compounds 1-2.

A. fumigatus strains were inoculated (1.0×10^6 spores mL^{-1}) into 1 L GMM[1] in a 2 L Erlenmeyer flask at 37 °C with shaking at 220 rpm. After 3.5 days, liquid fungal cultures including fungal tissue and media were frozen using a dry ice acetone bath, and lyophilized. The lyophilized residues were extracted with 500 mL of MeOH for 1 h with vigorous stirring. Extracts were filtered over cotton and evaporated on Celite-545 (Acros Organics) to dryness. The dry Celite was then loaded into a 25 gram solid phase cartridge (Isco) and subjected to large-scale reverse phase chromatography (see Chromatographic enrichment section below for details).

Analytical methods and equipment overview. (a) NMR spectroscopy: NMR spectroscopic instrumentation: a Bruker Avance^{III} HD (800 MHz ^{1}H reference frequency, 201 MHz for ^{13}C) equipped with a 5 mm CPTCL ^{1}H-^{13}C/^{15}N cryo probe. Non-gradient phase-cycled dqfCOSY spectra were acquired using the following parameters: 0.6 s acquisition time, 400-600 complex increments, 8, 16 or 32 scans per increment. Non-gradient HSQC, HMQC, and HMBC spectra were acquired with these parameters: 0.25 s acquisition time, 200-500 increments, 8-64 scans per increment. ^{1}H, ^{13}C-HMBC spectra were optimized for $J_{H,C}$ = 6 Hz. HSQC spectra were acquired with or without decoupling. NMR spectra were processed and baseline corrected using MestreLabs

MNOVA software packages. (b) Mass spectrometry: high-resolution LC-MS was performed on a Thermo Scientific-Dionex Ultimate3000 UHPLC system equipped with a diode array detector and connected to a Thermo Scientific Q Exactive Orbitrap operated in electrospray positive (ESI^+) or electrospray negative (ESI^-) ionization mode. Data acquisition and processing for the LC-HRMS was controlled by Thermo Scientific Xcalibur software. (c) Chromatography: flash chromatography was performed using a Teledyne ISCO CombiFlash system. For semi-preparative HPLC Agilent Zorbax Eclipse XDB-C18 or -C8 columns (25 cm x 10 mm, 5 µm particle diameter) were used. An Agilent Zorbax RRHD Eclipse XDB-C18 column (2.1 x 100 mm, 1.8 µm particle diameter) was used in the LC-HRMS *A. fumigatus* mutant profiling analysis.

Chromatographic enrichment of compounds 1-2. Methanol extracts derived from 1-2 L of *A. fumigatus* cultures were fractionated using large-scale reverse-phase flash chromatography on a Teledyne ISCO CombiFlash with a Teledyne C18 gold (100 gram) column with acetonitrile (organic phase) and 0.1 % acetic acid in water (aqueous phase) as solvents at a flow rate of 60 mL/min. A linear ramp from 0 % organic to 100 % organic over 30 min was used and fractions containing compounds of interest were collected, evaporated to dryness, and stored in 8 mL glass vials at -20 °C. Fractions containing compounds **1-2** were further purified via semi-preparative HPLC using an Agilent XDB C-18 or C-8 column (25 cm x 10 mm, 5 µm particle diameter) acetonitrile (organic phase) and 0.1 % acetic acid in water (aqueous phase) as solvents at a flow rate of 3.6 mL/min. A solvent gradient scheme was used, starting at 5 % organic for 3 min, followed by a linear increase to 100 % organic over 27 min, holding at 100 %

organic for 5 min, then decreasing back to 5 % organic for 0.1 min, and holding at 5 % organic for the final 4.9 min, for a total of 40 min.

***In vivo* stable isotope labeling experiments.** Deuterium- and ^{13}C-labeled amino acids were purchased from Cambridge Isotopes Inc. Each amino acid (labeled or unlabeled control) was added to 50 mL (2 mM) of GMM[1] media in 125 mL Erlenmeyer flasks. Each flask was inoculated with OE::*fsqA* or OE::*fsqA*-Δ*fsqF* (1.0×10^6 spores mL^{-1}) and grown for 4 days at 37 °C with shaking at 220 rpm. After 4 days, liquid fungal cultures including fungal tissue and media were frozen using a dry ice acetone bath and lyophilized. The lyophilized residues were extracted with MeOH (25 mL) for 1.5 h with vigorous stirring. Extracts were filtered over cotton, evaporated to dryness, and stored in 4 mL vials at -20 °C. Crude extracts were suspended in 0.5 mL of MeOH and centrifuged to remove insoluble materials, and the supernatant was subjected to LC-MS analysis.

Table E1. LC-HRMS data of reported compounds

Compound	HR-ESI(+/-) observed (*m/z*)	Ion	Calculated ion formula	Calculated *m/z*	Retention time [min]
1	251.1049	$[M-H]^-$	$C_{12}H_{15}N_2O_4^-$	251.1037	4.36
2	235.10	$[M-H]^-$	$C_{12}H_{15}N_2O_3^-$	235.1088	5.06

Table E2. Fungal strains and plasmids used in this study

Strain/plasmid	Description	Reference
***Aspergillus fumigatus* 293 background strains**		
Af293	Wild type	10
Af293.1	*pyrG1*	10
Af293.6	*pyrG1, argB1*	10
TJES1.18	*A.fargB::gpdA(p):: fsqA, pyrG1*	This study
TJES2.20	*ΔfsqA::A.ppyrG, argB1*	This study
TJES3.1	*A.fargB::gpdA(p):: fsqA, A.ppyrG*	This study
TJES4.10	*A.fargB::gpdA(p):: fsqA, ΔfsqF::A.ppyrG*	This study
TJES8.2	*ΔfsqA::A.ppyrG, A.fargB*	This study
TJES13.61	*A.fargB::gpdA(p):: fsqA, ΔfsqB::A.ppyrG*	This study
TJES14.68	*A.fargB::gpdA(p):: fsqA, ΔfsqC::A.ppyrG*	This study
TJES15.91	*A.fargB::gpdA(p):: fsqA, ΔfsqG::A.ppyrG*	This study
***Aspergillus flavus* NRRL3357 background strains**		
NRRL3357.5	*pyrG1*	11
TJW149.27	*Δku70::pyrG*	This study
TJES19.1	*ΔpyrG, Δku70*	This study
TJES 23.3	*A.ppyrG::gpdA(p)::imqA, Δku70*	This study
TJES 27.1	*ΔimqA::A.ppyrG, Δku70*	This study
Plasmids		
pJMP4	*A. fumigatus argB*	12
pJMP8.1	*A. nidulans* gpdA promoter (truncated to 1.5 kb) in pBluescript	13
pJMP9.1	*A. nidulans gpdA(p) + A parasiticus pyrG*	14
pJES1.2	*gpdA(p):: fsqA* in *pJMP8.1*	This study
pJES2.7	*A. fumigatus argB::gpdA(p):: fsqA*	This study
pJES13.2	C-terminal 6His tagged *fsqF* A-domain in pET30a vector	This study
pJW24	*A parasiticus pyrG* in pBluescript	15

pXX = plasmid, TXX = original transformant

Table E3. PCR primer sets used in this study

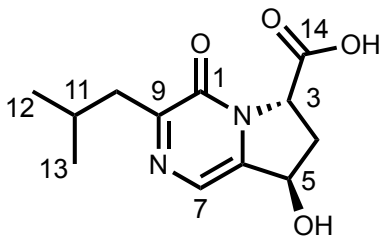
Name of the primer	Oligonucleotide sequence (5'-3')	Uses
Afu6g03380_FOR_Nprobe	CTTTCACGTGGACACTGCGC	Northern probe
Afu6g03380_REV_Nprobe	TACTGCTCCAACCAGCACCG	Northern probe
Afu6g03390_FOR_Nprobe	TCAGGGAGATATGGTGGCG	Northern probe
Afu6g03390_REV_Nprobe	CAACGCAGCAGGTAGTCACG	Northern probe
Afu6g03400_FOR_Nprobe	CTTCCAAGCCCAACAAGCC	Northern probe
Afu6g03400_REV_Nprobe	AATCTCGTAGGCCTCCAGCG	Northern probe
Afu6g03430_FOR_Nprobe	TGGCCTATCACACCAGTGGC	Northern probe
Afu6g03430_REV_Nprobe	GTGCAGCCTGAATCTCACGG	Northern probe
Afu6g03420_FOR_Nprobe	GAGCTCGACAATGGTGAGCG	Northern probe
Afu6g03420_REV_Nprobe	CAATCACATCGGCATGCGG	Northern probe
Afu6g03440_FOR_Nprobe	GATGGAACTTCGAGCCACCC	Northern probe
Afu6g03440_REV_Nprobe	GGCGATAATCTGCCAACGCC	Northern probe
Afu6g03450_FOR_Nprobe	CTACATCGCCTGCGATGTGG	Northern probe
Afu6g03450_REV_Nprobe	TAGCACACGCGCCAGATACC	Northern probe
Afu6g03460_FOR_Nprobe	GCGACTTCGCGACTCGGAAT	Northern probe
Afu6g03460_REV_Nprobe	CCATCACAAACTCGGTCCCG	Northern probe
Afu6g03470_FOR_Nprobe	AACTGCGCTCCAAAACCGCC	Northern probe
Afu6g03470_REV_Nprobe	ATCCACAAGGGCGATCTGGC	Northern probe
Afu6g03490_FOR_Nprobe	GTTCTCAGGGGATGTGACCG	Northern probe
Afu6g03490_REV_Nprobe	ACAAGTTCGCCCTTCGCTCCG	Northern probe
Afu6g03500_FOR_Nprobe	GGTGCTCAAGGAACAGAGGG	Northern probe
Afu6g03500_REV_Nprobe	GCCAAGAGGTCATTCTGCCC	Northern probe
Afu6g03510_FOR_Nprobe	AACACCCGATACCAGCTCGC	Northern probe
Afu6g03510_REV_Nprobe	ATGGGCCACCCATTGATGGC	Northern probe
Afu6g03520_FOR_Nprobe	ATGCCATCATCACCGGTGCC	Northern probe
Afu6g03520_REV_Nprobe	CGAGCATGGACAATAGCC	Northern probe
A.ppyrG_T7 FOR	CGTAATACGACTCACTATAGGG	Amplification of A.ppyrG from pJMP9.1
A.ppyrGR_Rev	ATTCGACAATCGGAGAGGCTGC	Amplification of A.ppyrG from pJMP9.1
Afu6g03430_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCG AATGCTTCAGCTGGAGTGTCTCC	Amplification of *fsqA* 3' flanking region
Afu6g03430_3'R_flank	TACAGCGACGACCAACGAGC	Amplification of *fsqA* 3' flanking region
Afu6g03430_5'F_flank	TAAGAGCGGAGACTGGTGGC	Amplification of *fsqA* 5' flanking region
Afu6g03430_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATT ACGTCTGCAAGGGTTTACGAGGG	Amplification of *fsqA* 5' flanking region
Afu6g03440_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCG	Amplification of *fsqB* 3'

	AATTCGGAACTCTGGAGGTTCCC	flanking region
Afu6g03440_3'R_flank	ACTGCGCGACAAATGCAGCC	Amplification of *fsqB* 3' flanking region
Afu6g03440_5'F_flank	GTCTCGTCACTTACCCTGCC	Amplification of *fsqB* 5' flanking region
Afu6g03440_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATT ACGAAAGAGACAGCCGGGATCCG	Amplification of *fsqB* 5' flanking region
Afu6g03450_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCG AATCTTGCTGCGGAAATCGAGCG	Amplification of *fsqC* 3' flanking region
Afu6g03450_3'R_flank	CACGGTAAAAGCCCAGTCCG	Amplification of *fsqC* 3' flanking region
Afu6g03450_5'F_flank	GATGTAGGCCACGAACTCGC	Amplification of *fsqC* 5' flanking region
Afu6g03450_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATT ACGAGGATGCCAAAAGCCCACCG	Amplification of *fsqC* 5' flanking region
Afu6g03480_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCG AATCGCGGGCATCTAGTATTCGG	Amplification of *fsqF* 3' flanking region
Afu6g03480_3'R_flank	ACTTGCGCAACCAGCTGTGC	Amplification of *fsqF* 3' flanking region
Afu6g03480_5'F_flank	GAATCTGAGCGCTTGTCGCG	Amplification of *fsqF* 5' flanking region
Afu6g03480_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATT ACGAAGAAAGGCGAAACGGAGCG	Amplification of *fsqF* 5' flanking region
Afu6g03490_3'F_flank	CTGTCGCTGCAGCCTCTCCGATTGTCG AATATGGACTCCAGTCAGGACCG	Amplification of *fsqG* 3' flanking region
Afu6g03490_3'R_flank	ATCCACCTCGTGGAGAAGCC	Amplification of *fsqG* 3' flanking region
Afu6g03490_5'F_flank	TGTGTCACGAAGGCAGTGCG	Amplification of *fsqG* 5' flanking region
Afu6g03490_5'R_flank	CCAATTCGCCCTATAGTGAGTCGTATT ACGGTGCCCATCGTCCAATACGG	Amplification of *fsqG* 5' flanking region
fqsF_Adomain_xp_5'F	AATAA**GCGGCCGC**ACCTGGAACACCG TGGTTGC	Amplification of *fsqF* A-domain **(NotI cut site)**
fqsF_Adomain_xp_3' R	ATTAGC**CTCGAG**TTCGTTGCCCCGAGT GTGC	Amplification of *fsqF* A-domain **(XhoI cut site)**
fsqD_xp_5' F	GTAGGCTA**GAATTC**TCACACCAGCCTC TCTTGGC	Amplification of *fsqD* ORF **(EcoRI cut site)**
fsqD_xp_3' R	CATTAT**AAGCTT**GAGAAAGGAGTAGC GGACCTCTTCCC	Amplification of *fsqD* ORF **(HindIII cut site)**
Afu6g03430_NcoI_FOR	CAGATG**CCATGG**ACGACAAGCATGGC C	Amplification of *fsqA* ORF **(NcoI cut site)**
Afu6g03430_NotI_REV	GCAACC**GCGGCCGC**AGACAGCGCGGT	Amplification of *fsqA*

	ATCACTG	ORF **(NotI cut site)**
flvku70F5	ACATCTCTTCCGTCAAAGGCGC	Amplification of *ku70* 5’ flanking region
flavku70R5	CGATATCAAGCTATCGATACCTCGACT CTGTGTTGAGAGTCGTAAGTCATGAAT TGCG	Amplification of *ku70* 5’ flanking region
flvku70F3	GTCGCTGCAGCCTCTCCGATTGTCGAA TGACAACGCTAGTATTGGTTACGAGAG ACAG	Amplification of *ku70* 3’ flanking region
flavku70R3	AGAATGGCTACGTCAACCTCCG	Amplification of *ku70* 3’ flanking region
Fku70IF	ATGAGGAAGAGGAGGAGACCG	Amplification of Δ*ku70* cassette
Fku70IR	CACTTTTCAATCGTGCGAGCCG	Amplification of Δ*ku70* cassette
pet28_fsqB_3’_fwd	TGGTGCCGCGCGGCAGCCATATGTC TATCCCTAACTCTTTCATCATTGT	Amplification of fsqB, pET28b+ 3’ destination
fsqB_5’_pet28_rev	CTCAGCTTCCTTTCGGGCTTTGTTACTA GTTGTCCTGTAGTAGTTCCACGG	Amplification of fsqB pET28b+ 5’ destination

Table E4. ^{1}H (800 MHz) and ^{13}C (201 MHz) NMR spectroscopic data for compound **1** in methanol-d_4.

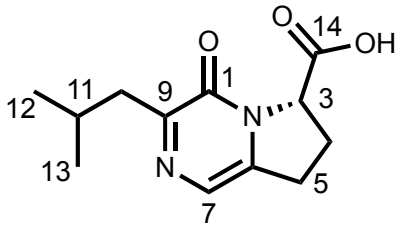
Chemical shifts were referenced to δ(C<u>H</u>D_2OD) = 3.31 ppm and δ(^{13}<u>C</u>HD_2OD_3) = 49.00.^{13}C chemical shifts were determined via HMBC, HSQC and direct observation ^{13}C spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. NOESY correlations were observed using a mixing time of 650 ms. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.



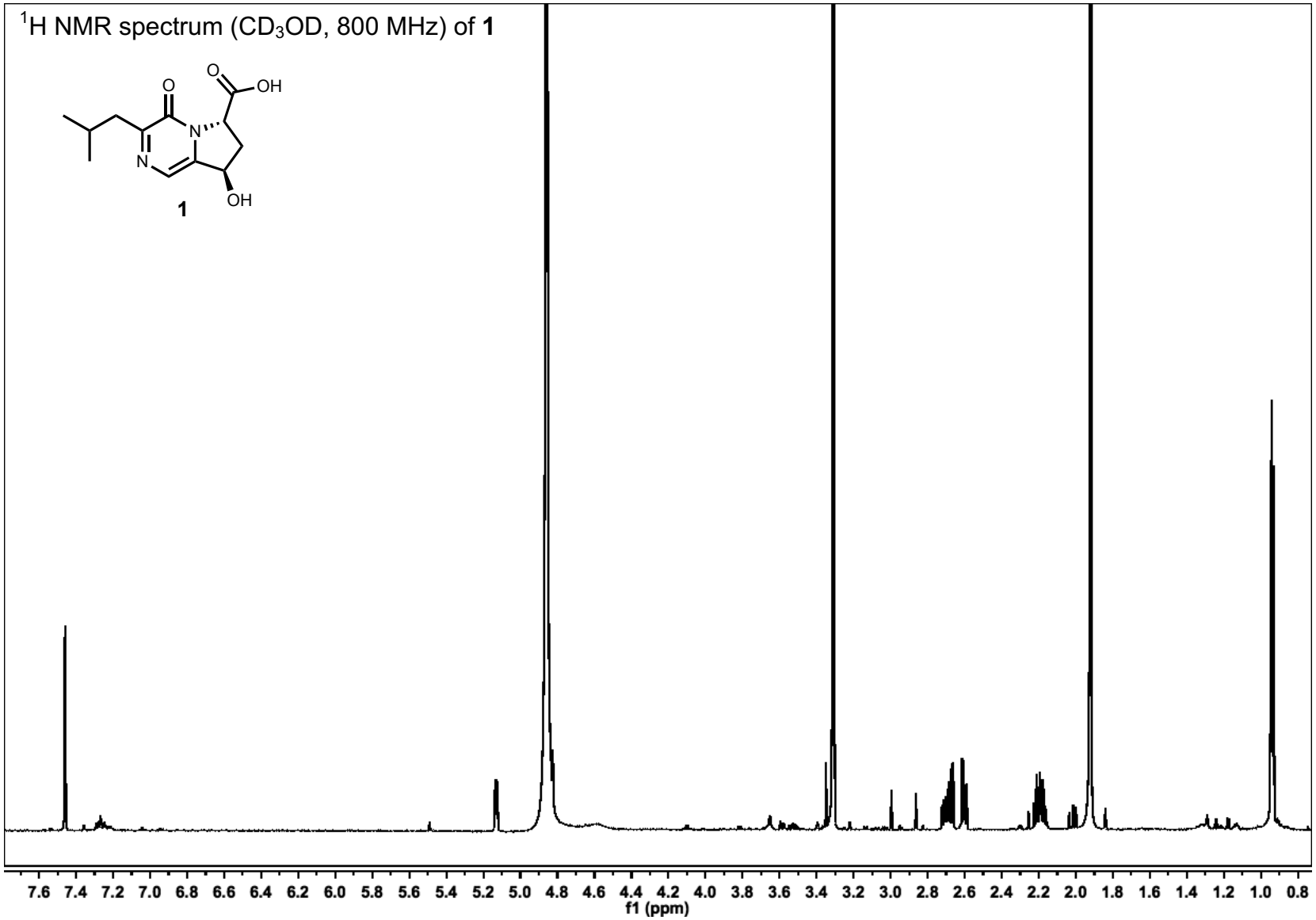
No.	δ_c	Proton	δH(J_{HH}[Hz])	HMBC	NOESY
1	156.46				
2					
3	64.95	3-H	4.83 ($J_{3,4a}$ = 3.8) ($J_{3,4b}$ = 8.5)	1,4,5,6,14	4
4	37.50	4-H_a	2.22 ($J_{4a,4b}$ = 13.5) ($J_{4a,3}$ = 3.8) ($J_{4a,5}$ = 4.0)	3,5,6,14	3,5
		4-H_b	2.72 ($J_{4b,4a}$ = 13.5) ($J_{4b,3}$ = 8.5) ($J_{4b,5}$ = 7.9)	3,5,6,14	3,5
5	72.43	5-H	5.15 ($J_{5,4a}$ = 4.0) ($J_{5,4b}$ = 7.9) ($J_{5,7}$ = 1.0)	3,7	4,7
6	144.39				
7	119.55	7-H	7.46 ($J_{7,5}$ = 1.0)	1w,5,6,9	5
8					
9	159.09				
10	42.37	10-H_a	2.60 ($J_{10a,10b}$ = 13.8) ($J_{10a,11}$ = 7.8)	1,9,11,12,13	11
		10-H_b	2.68 ($J_{10b,10a}$ = 13.8) ($J_{10b,11}$ = 6.4)	1,9,11,12,13	11
11	27.77	11-H	2.20 ($J_{11,10a}$ = 7.8) ($J_{11,10b}$ = 6.4) ($J_{11,12/13}$ = 7.0)	9,10,12,13	10
12	22.67	12-H_3	0.93 ($J_{12,11}$ = 7.0)	10,11,13	
13	22.67	13-H_3	0.93 ($J_{13,11}$ = 7.0)	10,11,12	
14	176.02				

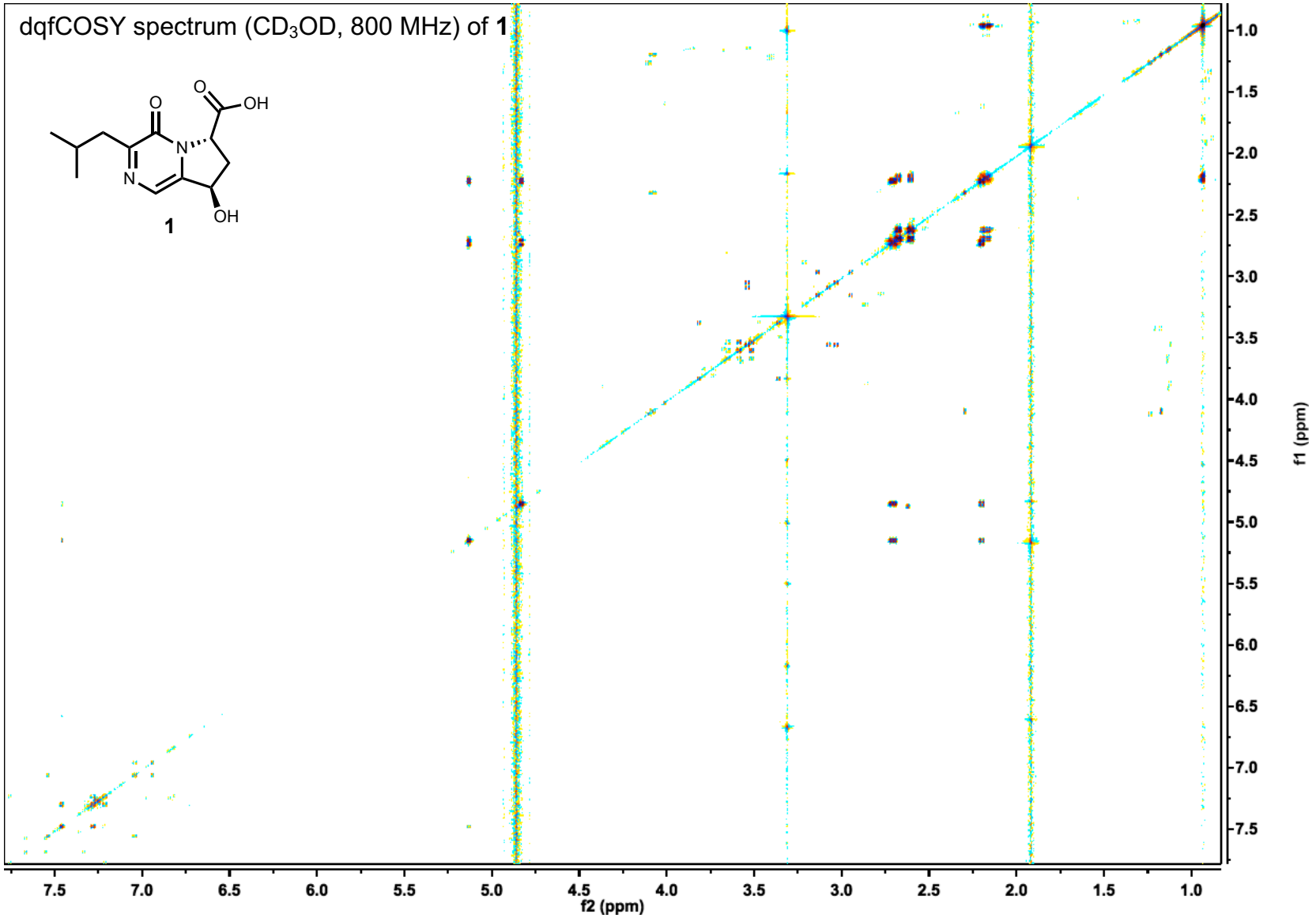
Table E5. ^{1}H (800 MHz) and ^{13}C (201 MHz) NMR spectroscopic data for compound **2** in methanol-d_4.

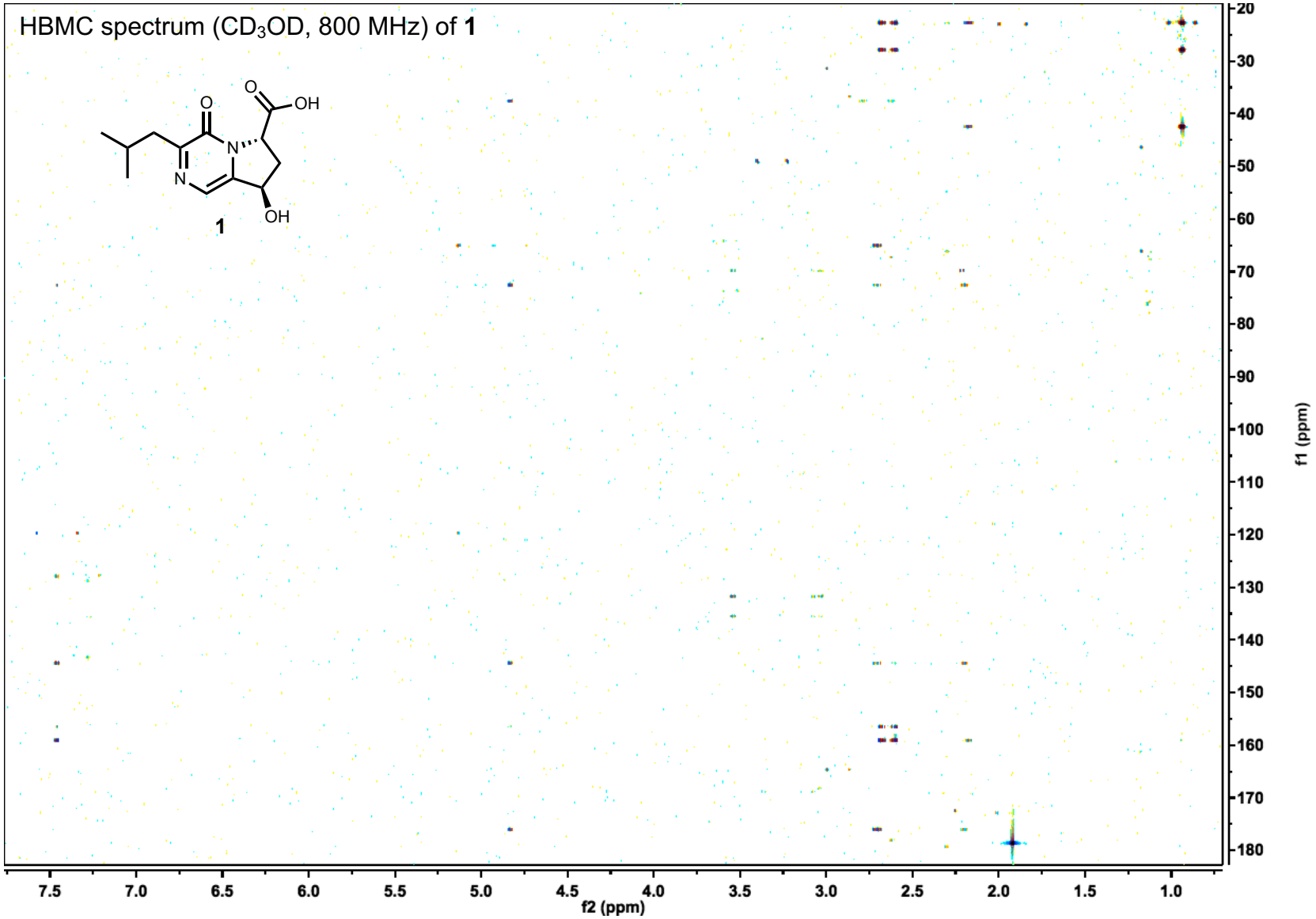
Chemical shifts were referenced to δ($C\underline{H}D_2OD$) = 3.31 ppm and δ($^{13}\underline{C}HD_2OD_3$) = 49.00.$^{13}C$ chemical shifts were determined via HMBC, HSQC and direct observation ^{13}C spectra. ^{1}H, ^{1}H-*J*-coupling constants were determined from the acquired ^{1}H or dqfCOSY spectra. HMBC correlations are from the proton(s) stated to the indicated ^{13}C atom.

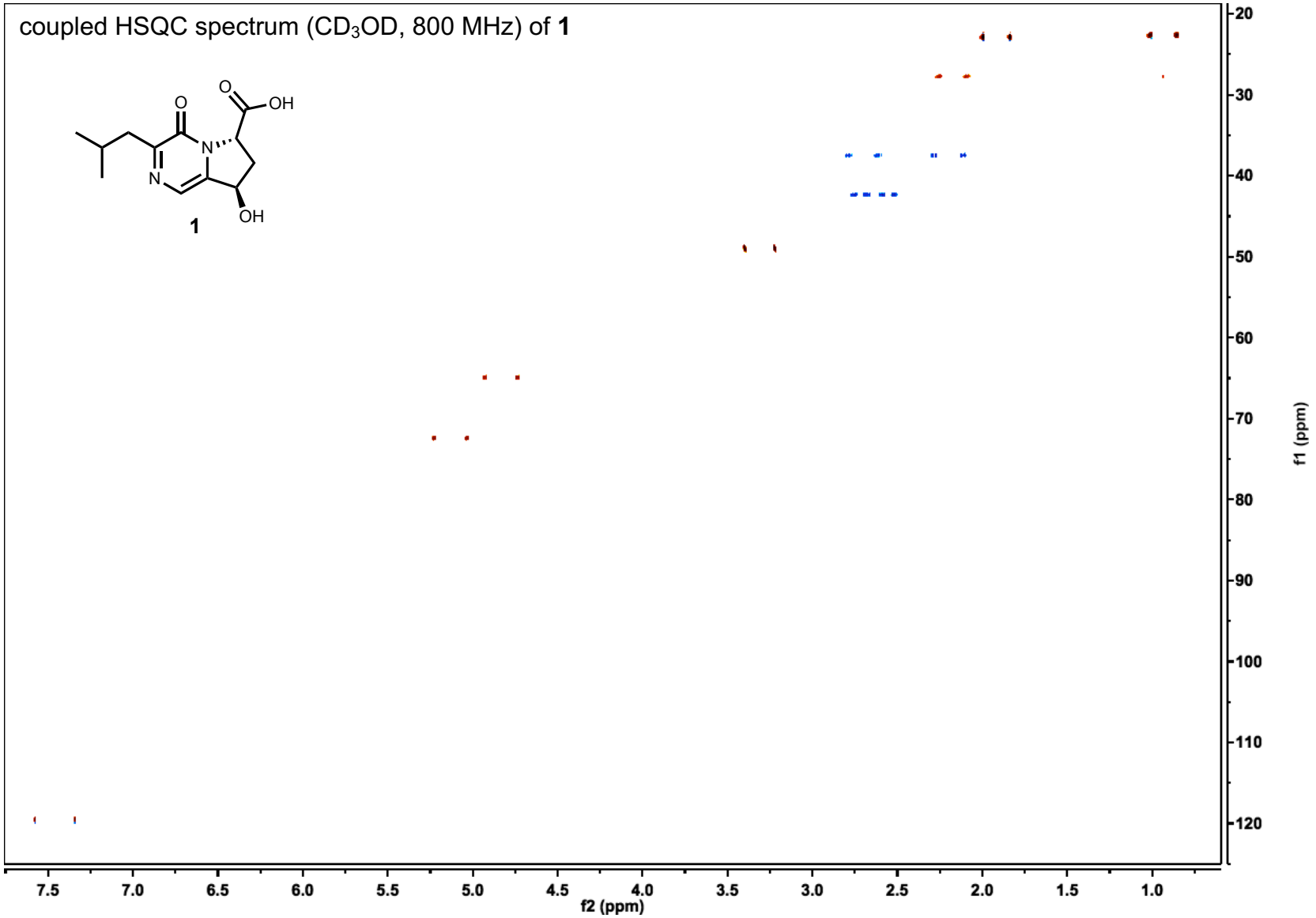


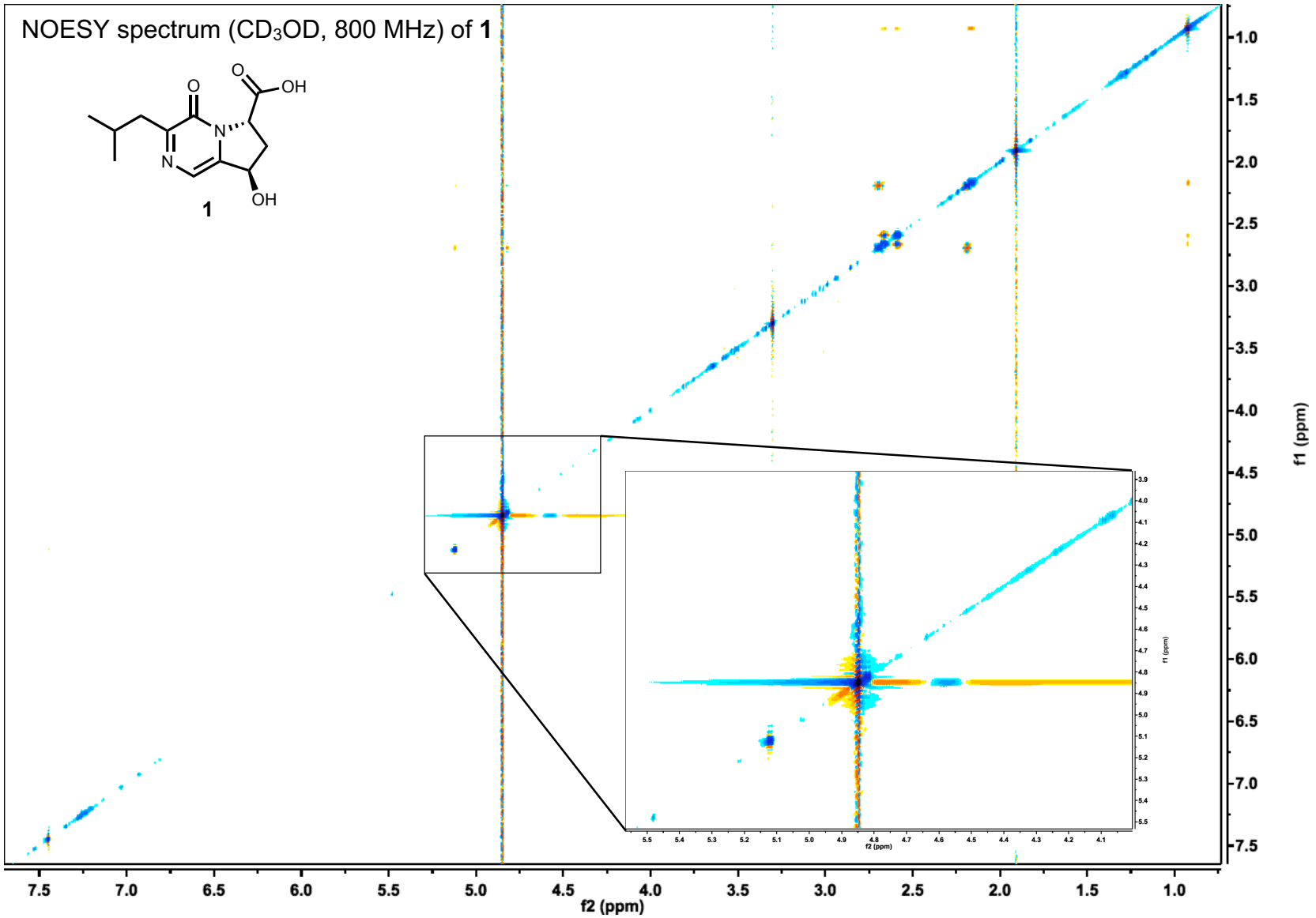
No.	δ_c	Proton	δH(J_{HH}[Hz])	HMBC
1	157.01			
2				
3		3-H	4.98(Broad)	
4	28.33	4-H_a	2.34 ($J_{4a,4b}$ = 15.6) ($J_{4a,3}$ = 13.0) ($J_{4a,5b}$ = 6.7)	
		4-H_b	2.48 ($J_{4a,4b}$ = 15.6) ($J_{4a,5}$ = 11.0)	
5	28.99	5-H_a	3.08 ($J_{5a,5b}$ = 17.0) ($J_{5a,4b}$ = 8.5)	
		5-H_b	3.18 ($J_{5b,5a}$ = 17.0) ($J_{5b,4a}$ = 6.7) ($J_{5b,4b}$ = 11.0)	
6	143.76			
7	118.99	7-H	7.29	6,9
8				
9	156.46			
10	42.23	10-H_a	2.15 ($J_{10a,10b}$ = 14.0) ($J_{10a,11}$ = 7.3)	9,11,12,13
		10-H_b	2.65 ($J_{10b,10a}$ = 14.0) ($J_{10b,11}$ = 6.9)	9,11,12,13
11	27.84	11-H	2.15 ($J_{11,10a}$ = 7.3) ($J_{11,10b}$ = 6.9) ($J_{11,12/13}$ = 6.5)	10,12,13
12	22.66	12-H_3	0.93 ($J_{12,11}$ = 6.5)	10,11,13
13	22.73	13-H_3	0.93 ($J_{13,11}$ = 6.5)	10,11,12
14				

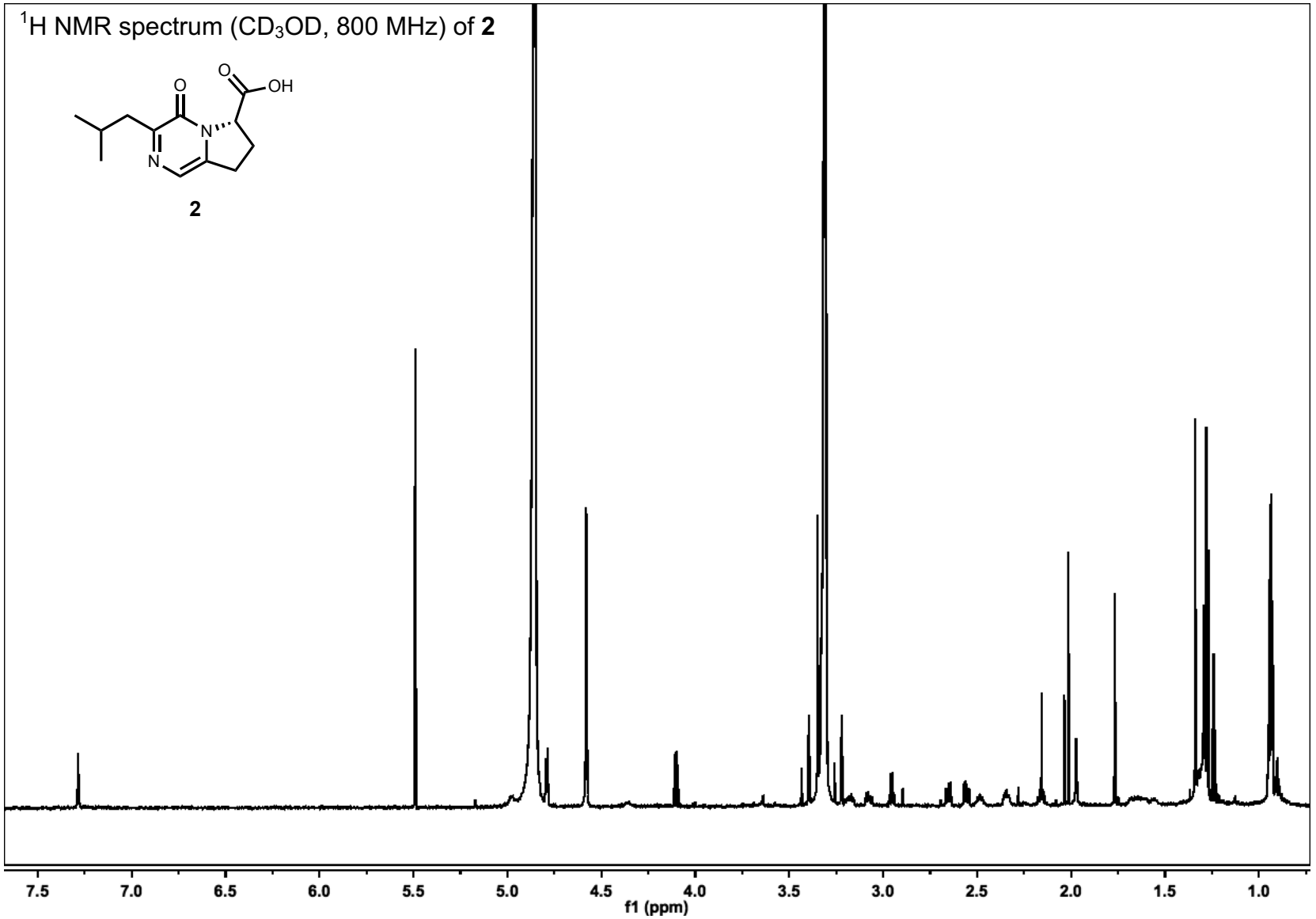

^{1}H NMR spectrum (CD_3OD, 800 MHz) of **1**
O
OH
N
N
OH
1
7.6 7.4 7.2 7.0 6.8 6.6 6.4 6.2 6.0 5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4 3.2 3.0 2.8 2.6 2.4 2.2 2.0 1.8 1.6 1.4 1.2 1.0 0.8
f1 (ppm)

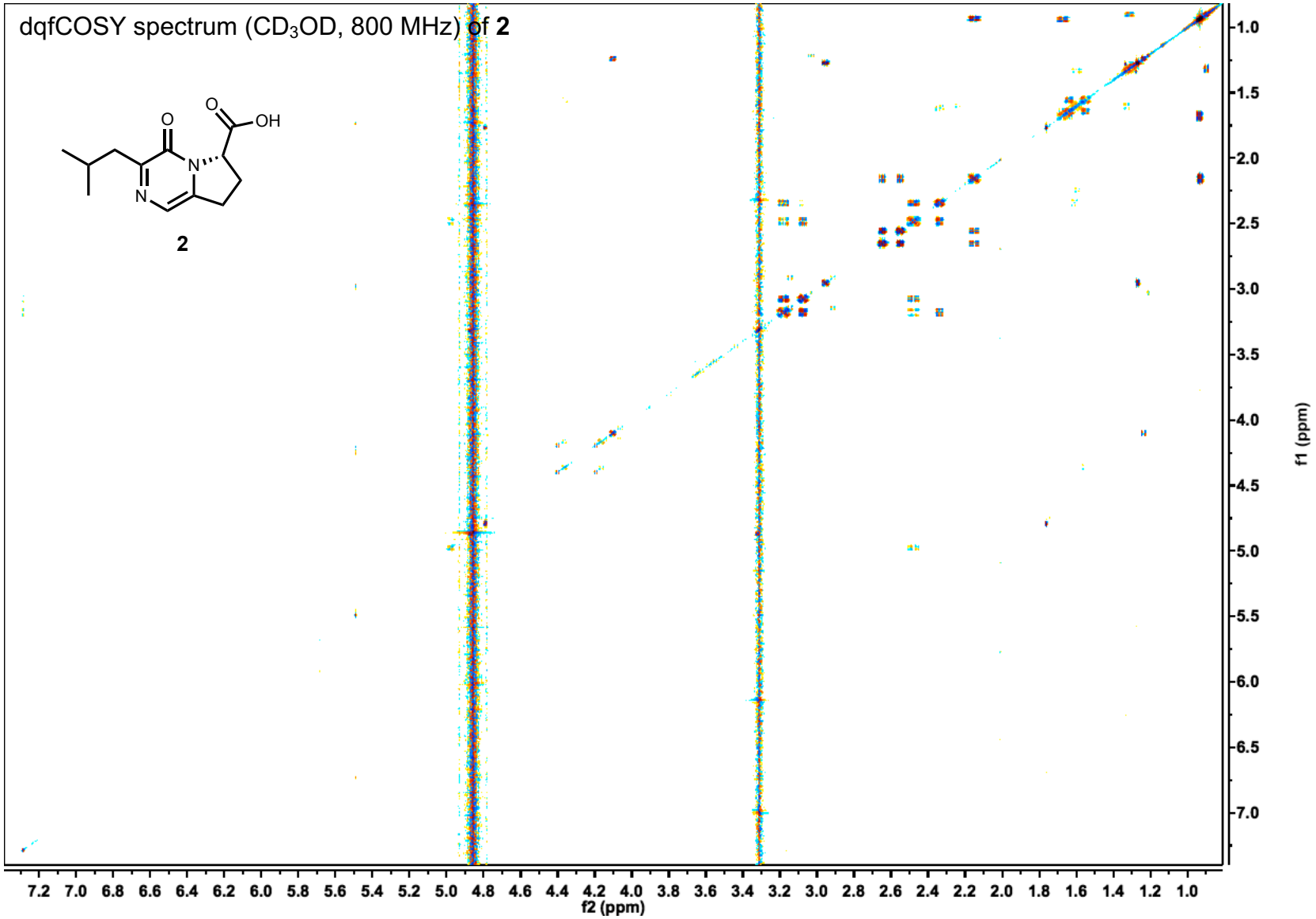

dqfCOSY spectrum (CD3OD, 800 MHz) of 1
O
OH
O
N
N
OH
1
f2 (ppm)
f1 (ppm)

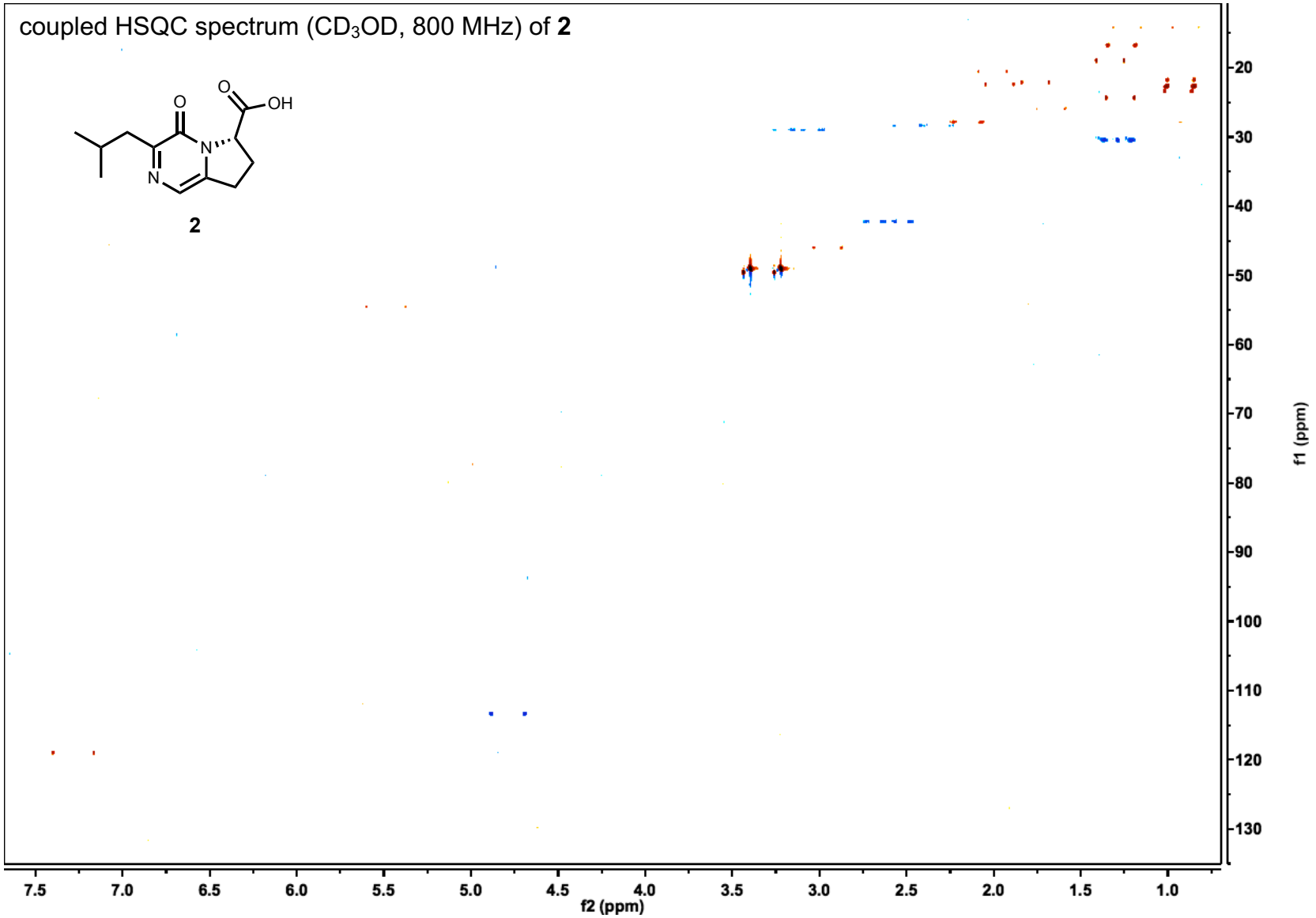

HBMC spectrum (CD3OD, 800 MHz) of 1
O
OH
N
N
O
OH
1
f2 (ppm)
f1 (ppm)

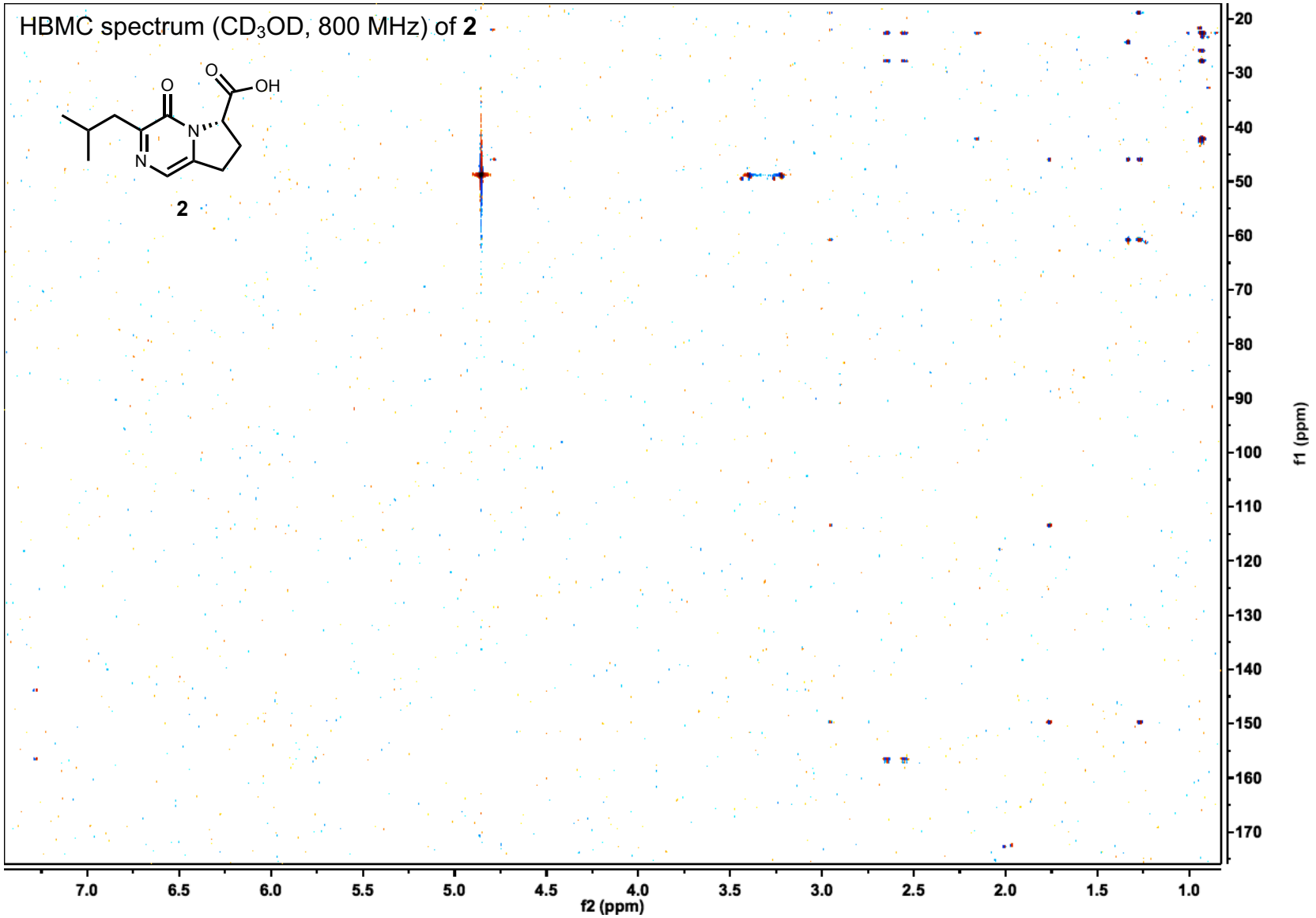

coupled HSQC spectrum (CD_3OD, 800 MHz) of **1**
O
OH
N
N
O
OH
1
f1 (ppm)
f2 (ppm)


NOESY spectrum (CD3OD, 800 MHz) of 1
O
OH
N
N
O
OH
1
f2 (ppm)
f1 (ppm)


^{1}H NMR spectrum (CD_3OD, 800 MHz) of **2**
O
O
OH
N
N
2
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)


dqfCOSY spectrum (CD_3OD, 800 MHz) of **2**
O
O
OH
N
N
2
f1 (ppm)
f2 (ppm)


coupled HSQC spectrum (CD_3OD, 800 MHz) of **2**
O
O
OH
N
N
2
f1 (ppm)
20
30
40
50
60
70
80
90
100
110
120
130
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f2 (ppm)


HBMC spectrum (CD3OD, 800 MHz) of 2
O
O
OH
N
N
2
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f2 (ppm)
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
f1 (ppm)

REFERENCES

1. Shimizu, K. & Keller, N. P. Genetic involvement of a cAMP-dependent protein kinase in a G protein signaling pathway regulating morphological and chemical transitions in Aspergillus nidulans. *Genetics* **157**, 591-600, (2001).

2. Sekonyela, R. *et al.* RsmA regulates *Aspergillus fumigatus* gliotoxin cluster metabolites including cyclo(L-Phe-L-Ser), a potential new diagnostic marker for invasive aspergillosis. *PLoS ONE* **8**, e62591, (2013).

3. Palmer, J. M. *et al.* Loss of CclA, required for histone 3 lysine 4 methylation, decreases growth but increases secondary metabolite production in Aspergillus fumigatus. *PeerJ* **1**, e4, (2013).

4. Yu, J. H. *et al.* Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi. *Fungal Genet Biol* **41**, 973-981, (2004).

5. Kalb, D., Lackner, G. & Hoffmeister, D. Functional and phylogenetic divergence of fungal adenylate-forming reductases. *Appl Environ Microbiol* **80**, 6175-6183, (2014).

6. Schneider, P., Bouhired, S. & Hoffmeister, D. Characterization of the atromentin biosynthesis genes and enzymes in the homobasidiomycete Tapinella panuoides. *Fungal Genet Biol* **45**, 1487-1496, (2008).

7. Bok, J. W. *et al.* Chromatin-level regulation of biosynthetic gene clusters. *Nat Chem Biol* **5**, 462-464, (2009).

8. Carpenter, J. F. An improved synthesis of 5,6-diacetoxy-N-methylindole and of epinine. *J Org Chem* **58**, 1607-1609, (1993).

9. Clauser, K. R., Baker, P. & Burlingame, A. L. Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem* **71,** 2871–82, (1999).

10. Xue, T., Nguyen, C. K., Romans, A., Kontoyiannis, D. P. & May, G. S. Isogenic auxotrophic mutant strains in the Aspergillus fumigatus genome reference strain AF293. *Arch Microbiol* **182,** 346-353, (2004).

11. He, Z., Price, M.S., OBrian, G.R., Georgianna, D.R. & Payne, G.A. Improved protocols for functional analysis in the pathogenic fungus Aspergillus flavus. *BMC Microbiol* **7**, 104, (2007).

12. Palmer, J. M. *et al.* Loss of CclA, required for histone 3 lysine 4 methylation, decreases growth but increases secondary metabolite production in Aspergillus fumigatus. *PeerJ* **1**, e4, (2013).

13. Sekonyela, R. *et al.* RsmA regulates Aspergillus fumigatus gliotoxin cluster metabolites including cyclo(L-Phe-L-Ser), a potential new diagnostic marker for invasive aspergillosis. *PLoS One* **8**, e62591, (2013).

14. Lim, F.Y., *et al.* Genome-Based Cluster Deletion Reveals an Endocrin Biosynthetic Pathway in Aspergillus fumigatus. *Appl Environ Microbiol* **78**, 4117-4125, (2012).

15. Calvo, A.M., Bok, J.W., Brooks, W. & Keller, N.P. LeA Is Required for Toxin and Sclerotial Production in Aspergillus parasiticus. *Appl Environ Microbiol* **70**, 4733-4739, (2004)